Measurement of Angular Observables of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ Decays and the Upgrade of LHCb

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Federica Lionetto

aus

Italien

Promotionskommission

Prof. Dr. Nicola Serra (Vorsitz)

PD Dr. Olaf Steinkamp

Dr. Katharina Müller

Zürich, 2018

Abstract

The LHCb experiment at the Large Hadron Collider is tailored to search for indirect evidences of physics beyond the Standard Model, mainly by studying *CP* violation and rare decays of charm and bottom hadrons. Despite the great success of LHCb, several key measurements will be limited by statistical uncertainties even after seven years of data taking. In order to increase the annual event yields, the LHCb detector will undergo a major upgrade during the Long Shutdown II, which will require the replacement of several subdetectors. The tracking system, in particular, will be completely redesigned in order to deal with the higher instantaneous luminosity and guarantee a full read out at the proton-proton collision rate of 40 MHz.

The first part of this thesis describes the design and testing of a new tracking detector, the Upstream Tracker. As its predecessor, the Upstream Tracker consists of silicon microstrip sensors. Its design involves two key elements that require dedicated measurements on prototype sensors before production: the adoption of an embedded pitch adapter to connect the silicon microstrips to the readout electronics and the configuration of the biasing scheme. Tests on prototype sensors were performed with beams of particles at the Super Proton Synchrotron in the past years and allowed to characterise the long-term performances of the sensors in terms of charge collection efficiency, cross talk, and radiation hardness. These studies resulted in the final design of the outermost sensors of the Upstream Tracker.

One of the most promising areas of research of flavour physics is the study of rare decays of bottom hadrons and, in particular, of those proceeding through a $b \to s\ell^+\ell^-$ transition, with $\ell = e, \mu$. These decays are particularly powerful tools to probe New Physics contributions, since the latter can result in sizeable deviations from the Standard Model predictions. In recent years, hints of deviations were observed in the angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$ decays, *e.g.* the P_5' anomaly, as well as in branching fraction ratios between muons and electrons, like R_K and R_{K^*}. These observations strongly increased the interest toward New Physics models in which lepton flavour universality is violated. Breaking of lepton flavour universality through New Physics effects can be measured by comparing the values of the $A_{\rm FB}$, S, and P' observables of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays.

The second part of this thesis describes the first measurement of the $\Delta A_{\rm FB} = A_{\rm FB}^e - A_{\rm FB}^\mu$ and $\Delta S_i = S_i^e - S_i^\mu$ observables, with $i = 4, 5, 7, 8$, using the dataset collected by LHCb from 2010 to 2012. These observables are expected to be compatible with zero in the Standard Model, but can significantly deviate from this value in the event of New Physics. The analysis described in this thesis is performed using a counting method procedure, which is based on the symmetry properties of the *CP*-average differential decay width of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decay. This approach is very robust, even with limited statistics, and is independent of the angular distribution of background contributions. The measurement is performed in two bins of q^2 (that is, of the dilepton invariant mass squared), in the low q^2 region from $0.1\,{\rm GeV}^2/c^4$ to $1.1\,{\rm GeV}^2/c^4$, and in the central q^2 region from $1.1\,{\rm GeV}^2/c^4$ to $7.0\,{\rm GeV}^2/c^4$. The results are compatible with the Standard Model predictions and are dominated by statistical uncertainties. The analysis presented here shows the capabilities of the LHCb experiment in measuring the angular observables of $B^0 \to K^{*0}\ \ell^+\ell^-$ decays and provides a baseline for upcoming measurements, which will include the dataset collected since 2012.

Zusammenfassung

Das LHCb-Experiment am Large Hadron Collider ist auf die Suche nach indirekten Hinweisen auf Physik jenseits des Standardmodells spezialisiert, hauptsächlich durch die Untersuchung von *CP*-Verletzung und seltenen Zerfällen von Charm und Bottom Hadronen. Trotz des grossen Erfolges von LHCb werden einige Schlüsselmessungen auch nach eine Messperiode von 7 Jahren immer noch durch statistische Unsicherheiten eingeschränkt sein. Um die jährliche Ereignisausbeute zu erhöhen, wird der LHCb-Detektor während des Long Shutdown II verbessert und erweitert, was den Austausch mehrerer Subdetektoren erfordert. Insbesondere die Spurkammern werden komplett überarbeitet, um der höheren momentanen Leuchtkraft gerecht zu werden und ein vollständiges Auslesen des Detektors bei der Proton-Proton-Kollisionsrate von 40 MHz zu gewährleisten.

Der erste Teil dieser Arbeit beschreibt das Design und die Untersuchung eines neuen Spur-Detektors, des Upstream Tracker. Der Upstream Tracker besteht wie sein Vorgänger aus Silizium-Mikrostreifen-Sensoren. Sein Design beinhaltet zwei Schlüsselelemente, die spezielle Messungen an Prototypen-Sensoren vor der Produktion erfordern: die Einführung eines Embedded Pitch Adapters zur Verbindung der Silizium-Mikrostreifen mit der Ausleseelektronik und die Konfiguration des Biasingsschemas. Dedizierte Messungen an Prototyp-Sensoren wurden in den letzten Jahren mit Teilchenstrahlen am Super Proton Synchrotron durchgeführt. Diese Messungen zielten darauf ab, die Langzeitleistungen der Sensoren hinsichtlich Effizienz der Ladungserfassung, Übersprechen und Strahlungshärte zu charakterisieren und führten zum endgültigen Design der Sensoren.

Eines der vielversprechendsten Forschungsgebiete der Physik mit schwere Quarks ist die Erforschung der seltenen Zerfälle von Bottom Hadronen und insbesondere derjenigen, die durch einen $b \to s\ell^+\ell^-$-Übergang stattfinden ($\ell = e, \mu$). Diese Zerfälle sind besonders wichtig, um Beiträge von Neuer Physik zu untersuchen, da letztere zu erheblichen Abweichungen von den Vorhersagen des Standardmodells führen können. In den letzten Jahren wurden Hinweise auf Abweichungen in der Winkelverteilung von $B^0 \to K^{*0}\mu^+\mu^-$-Zerfällen beobachtet, z.B. in der P_5'-Anomalie, sowie in Quotienten von Verzweigungsverhältnissen zwischen Myonen und Elektronen, wie R_K und R_{K^*}. Diese Beobachtungen haben das Interesse an Modellen der Neuen Physik, in denen die Universalität des Lepton-Flavour verletzt wird, stark erhöht. Eine Verletzung der Universalität des Lepton-Flavour durch Beiträge Neuer Physik kann gemessen werden, indem man die Werte der $A_{\rm FB}$-, S- und P'-Beobachtungen der Zerfälle $B^0 \to K^{*0}\mu^+\mu^-$ und $B^0 \to K^{*0}e^+e^-$ vergleicht.

Der zweite Teil dieser Arbeit beschreibt die erste Messung des $\Delta A_{\rm FB} = A^e_{\rm FB} - A^\mu_{\rm FB}$ und $\Delta S_i = S^e_i - S^\mu_i$ Observabeln, mit $i = 4, 5, 7, 8$, unter Verwendung des Datensatzes, der von LHCb 2010-2012 gesammelt wurde. Diese Observabeln sind im Standardmodell mit Null kompatibel, können aber im Falle Neuer Physik deutlich von diesem Wert abweichen. Die in dieser Arbeit beschriebene Analyse wird mit Hilfe einer Zählmethode durchgeführt, die auf den Symmetrieeigenschaften der über *CP*-gemittelten differentiellen Zerfallsbreite des $B^0 \to K^{*0}\ \ell^+\ell^-$ Zerfalls basiert. Dieser Ansatz ist auch bei kleine Statistik sehr stabil und unabhängig von der Winkelverteilung der Untergrundbeiträge. Die Messung wird in zwei Regionen von q^2, der quadratischen invarianten Masse des Leptonpaares, durchgeführt, in der unteren q^2-Region von $0.1\,{\rm GeV}^2/c^4$ bis $1.1\,{\rm GeV}^2/c^4$, und in der zentralen q^2-Region von $1.1\,{\rm GeV}^2/c^4$ bis $7.0\,{\rm GeV}^2/c^4$. Die Ergebnisse sind kompatibel mit den Vorhersagen des Standardmodells und werden nach wie vor von der statistischen Unsicherheit dominiert. Die hier vorgestellte Analyse zeigt das Potential des LHCb-Experiments für Messungen der Winkelverteilung von $B^0K^{*0}\ \ell^+\ell^-$-Zerfällen und liefert eine Grundlage für zukünftige Messungen.

Alla mia famiglia,
che mi è sempre stata accanto
e continua a condividere con me i piccoli e i grandi momenti di ogni giorno.

Contents

List of Figures

List of Tables

Abstract

The LHCb experiment at the Large Hadron Collider is tailored to search for indirect evidences of physics beyond the Standard Model, mainly by studying *CP* violation and rare decays of charm and bottom hadrons. Despite the great success of LHCb, several key measurements will be limited by statistical uncertainties even after seven years of data taking. In order to increase the annual event yields, the LHCb detector will undergo a major upgrade during the Long Shutdown II, which will require the replacement of several subdetectors. The tracking system, in particular, will be completely redesigned in order to deal with the higher instantaneous luminosity and guarantee a full read out at the proton-proton collision rate of 40 MHz.

The first part of this thesis describes the design and testing of a new tracking detector, the Upstream Tracker. As its predecessor, the Upstream Tracker consists of silicon microstrip sensors. Its design involves two key elements that require dedicated measurements on prototype sensors before production: the adoption of an embedded pitch adapter to connect the silicon microstrips to the readout electronics and the configuration of the biasing scheme. Tests on prototype sensors were performed with beams of particles at the Super Proton Synchrotron in the past years and allowed to characterise the long-term performances of the sensors in terms of charge collection efficiency, cross talk, and radiation hardness. These studies resulted in the final design of the outermost sensors of the Upstream Tracker.

One of the most promising areas of research of flavour physics is the study of rare decays of bottom hadrons and, in particular, of those proceeding through a $b \to s\ell^+\ell^-$ transition, with $\ell = e, \mu$. These decays are particularly powerful tools to probe New Physics contributions, since the latter can result in sizeable deviations from the Standard Model predictions. In recent years, hints of deviations were observed in the angular distribution of $B^0 \to K^{*0}\mu^+\mu^-$ decays, *e.g.* the P_5' anomaly, as well as in branching fraction ratios between muons and electrons, like R_K and R_{K^*}. These observations strongly increased the interest toward New Physics models in which lepton flavour universality is violated. Breaking of lepton flavour universality through New Physics effects can be measured by comparing the values of the $A_{\rm FB}$, S, and P' observables of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays.

The second part of this thesis describes the first measurement of the $\Delta A_{\rm FB} = A_{\rm FB}^e - A_{\rm FB}^\mu$ and $\Delta S_i = S_i^e - S_i^\mu$ observables, with $i = 4, 5, 7, 8$, using the dataset collected by LHCb from 2010 to 2012. These observables are expected to be compatible with zero in the Standard Model, but can significantly deviate from this value in the event of New Physics. The analysis described in this thesis is performed using a counting method procedure, which is based on the symmetry properties of the *CP*-average differential decay width of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decay. This approach is very robust, even with limited statistics, and is independent of the angular distribution of background contributions. The measurement is performed in two bins of q^2 (that is, of the dilepton invariant mass squared), in the low q^2 region from $0.1\,{\rm GeV}^2/c^4$ to $1.1\,{\rm GeV}^2/c^4$, and in the central q^2 region from $1.1\,{\rm GeV}^2/c^4$ to $7.0\,{\rm GeV}^2/c^4$. The results are compatible with the Standard Model predictions and are dominated by statistical uncertainties. The analysis presented here shows the capabilities of the LHCb experiment in measuring the angular observables of $B^0 \to K^{*0}\ \ell^+\ell^-$ decays and provides a baseline for upcoming measurements, which will include the dataset collected since 2012.

Zusammenfassung

Das LHCb-Experiment am Large Hadron Collider ist auf die Suche nach indirekten Hinweisen auf Physik jenseits des Standardmodells spezialisiert, hauptsächlich durch die Untersuchung von *CP*-Verletzung und seltenen Zerfällen von Charm und Bottom Hadronen. Trotz des grossen Erfolges von LHCb werden einige Schlüsselmessungen auch nach eine Messperiode von 7 Jahren immer noch durch statistische Unsicherheiten eingeschränkt sein. Um die jährliche Ereignisausbeute zu erhöhen, wird der LHCb-Detektor während des Long Shutdown II verbessert und erweitert, was den Austausch mehrerer Subdetektoren erfordert. Insbesondere die Spurkammern werden komplett überarbeitet, um der höheren momentanen Leuchtkraft gerecht zu werden und ein vollständiges Auslesen des Detektors bei der Proton-Proton-Kollisionsrate von 40 MHz zu gewährleisten.

Der erste Teil dieser Arbeit beschreibt das Design und die Untersuchung eines neuen Spur-Detektors, des Upstream Tracker. Der Upstream Tracker besteht wie sein Vorgänger aus Silizium-Mikrostreifen-Sensoren. Sein Design beinhaltet zwei Schlüsselelemente, die spezielle Messungen an Prototypen-Sensoren vor der Produktion erfordern: die Einführung eines Embedded Pitch Adapters zur Verbindung der Silizium-Mikrostreifen mit der Ausleseelektronik und die Konfiguration des Biasingsschemas. Dedizierte Messungen an Prototyp-Sensoren wurden in den letzten Jahren mit Teilchenstrahlen am Super Proton Synchrotron durchgeführt. Diese Messungen zielten darauf ab, die Langzeitleistungen der Sensoren hinsichtlich Effizienz der Ladungserfassung, Übersprechen und Strahlungshärte zu charakterisieren und führten zum endgültigen Design der Sensoren.

Eines der vielversprechendsten Forschungsgebiete der Physik mit schwere Quarks ist die Erforschung der seltenen Zerfälle von Bottom Hadronen und insbesondere derjenigen, die durch einen $b \to s\ell^+\ell^-$-Übergang stattfinden ($\ell = e, \mu$). Diese Zerfälle sind besonders wichtig, um Beiträge von Neuer Physik zu untersuchen, da letztere zu erheblichen Abweichungen von den Vorhersagen des Standardmodells führen können. In den letzten Jahren wurden Hinweise auf Abweichungen in der Winkelverteilung von $B^0 \to K^{*0}\mu^+\mu^-$-Zerfällen beobachtet, z.B. in der P_5'-Anomalie, sowie in Quotienten von Verzweigungsverhältnissen zwischen Myonen und Elektronen, wie R_K und R_{K^*}. Diese Beobachtungen haben das Interesse an Modellen der Neuen Physik, in denen die Universalität des Lepton-Flavour verletzt wird, stark erhöht. Eine Verletzung der Universalität des Lepton-Flavour durch Beiträge Neuer Physik kann gemessen werden, indem man die Werte der $A_{\rm FB}$-, S- und P'-Beobachtungen der Zerfälle $B^0 \to K^{*0}\mu^+\mu^-$ und $B^0 \to K^{*0}e^+e^-$ vergleicht.

Der zweite Teil dieser Arbeit beschreibt die erste Messung des $\Delta A_{\rm FB} = A^e_{\rm FB} - A^\mu_{\rm FB}$ und $\Delta S_i = S^e_i - S^\mu_i$ Observabeln, mit $i = 4, 5, 7, 8$, unter Verwendung des Datensatzes, der von LHCb 2010-2012 gesammelt wurde. Diese Observabeln sind im Standardmodell mit Null kompatibel, können aber im Falle Neuer Physik deutlich von diesem Wert abweichen. Die in dieser Arbeit beschriebene Analyse wird mit Hilfe einer Zählmethode durchgeführt, die auf den Symmetrieeigenschaften der über *CP*-gemittelten differentiellen Zerfallsbreite des $B^0 \to K^{*0}\ \ell^+\ell^-$ Zerfalls basiert. Dieser Ansatz ist auch bei kleine Statistik sehr stabil und unabhängig von der Winkelverteilung der Untergrundbeiträge. Die Messung wird in zwei Regionen von q^2, der quadratischen invarianten Masse des Leptonpaares, durchgeführt, in der unteren q^2-Region von $0.1\,{\rm GeV}^2/c^4$ bis $1.1\,{\rm GeV}^2/c^4$, und in der zentralen q^2-Region von $1.1\,{\rm GeV}^2/c^4$ bis $7.0\,{\rm GeV}^2/c^4$. Die Ergebnisse sind kompatibel mit den Vorhersagen des Standardmodells und werden nach wie vor von der statistischen Unsicherheit dominiert. Die hier vorgestellte Analyse zeigt das Potential des LHCb-Experiments für Messungen der Winkelverteilung von $B^0K^{*0}\ \ell^+\ell^-$-Zerfällen und liefert eine Grundlage für zukünftige Messungen.

Introduction

"The Standard Model of particle physics says that the universe consists of a very small number of particles, 12, and a very small number of forces, four. If we're correct about those 12 particles and those four forces and understand how they interact, properly, we have the recipe for baking up a universe." – Robert J. Sawyer

Despite the simplicity of this sentence, unveiling the laws that govern the interactions among elementary particles has kept several generations of physicists busy for decades. The Standard Model has been extensively tested by experiments, which confirmed all its predictions up to the $\mathcal{O}(1)$ TeV scale. However, some open questions still remain, which suggest the need for a more fundamental theory involving new particles and interactions at higher energies than probed so far. The theoretical framework of the Standard Model and some of its potential New Physics extensions are discussed in *Chapter 1*. After a general introduction, the chapter focuses on rare decays proceeding through a $b \to s\ell^+\ell^-$ transition. Two different approaches to search for New Physics effects in this transition are presented by briefly reviewing some of the key measurements previously performed at LHCb. The phenomenological interpretation of the current experimental results is also discussed. The last part of the chapter is devoted to introduce the angular asymmetries between decays to electrons and muons, referred to as ΔS_i, Q_i, and D_i observables, which can provide fundamental insights on the existence of lepton flavour universality breaking.

The LHCb experiment at the Large Hadron Collider is leading the exploration of indirect searches of New Physics in rare decays of charm and bottom hadrons. These decays have typical branching fractions of the order of 10^{-6} and are still partially unexplored. Only recently, experiments started to accumulate samples large enough to probe them with sufficient sensitivity. In order to further improve its data taking capabilities, the LHCb detector will undergo a major upgrade during the Long Shutdown II. Most of the subdetectors will be replaced in order to allow a higher instantaneous luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and an increased readout rate of 40 MHz. The first part of this thesis focuses on the upgrade and, in particular, on the design and testing of the Upstream Tracker, the silicon microstrip detector that will replace the Tracker Turicensis. In order to finalise the design of the detector, several test beam measurements were performed on prototype sensors in the past years. These measurements aimed at assessing the long-term performances of the silicon sensors of the

Upstream Tracker in terms of charge collection efficiency, cross talk, and radiation hardness. Two main aspects of the sensor design were investigated. The first was the adoption of an embedded pitch adapter to connect the silicon microstrips to the readout electronics; the second was the design of the biasing scheme and its potential impact on the sensor performances. A description of the LHCb detector and its performance can be found in *Chapter 2*, while *Chapter 3* gives an overview of the upgrade programme, with a particular focus directed to the Upstream Tracker. *Chapter 4* contains a brief description of the working principle of silicon detectors and introduces the key concepts that will be used extensively in the following chapter, while the test beam measurements performed on the prototype sensors are described in detail in *Chapter 5*.

The second part of this thesis describes the first measurement of the $A_{\rm FB}$, S_4, S_5, S_7, and S_8 angular observables of the $B^0 \to K^{*0}e^+e^-$ decay, as well as the measurement of the differences of these observables between electrons and muons. The measurement is based on the full Run I dataset collected by LHCb and is performed in two bins of q^2 (that is, of the dilepton invariant mass squared): in the low q^2 region from $0.1\,{\rm GeV}^2/c^4$ to $1.1\,{\rm GeV}^2/c^4$ and in the central q^2 region from $1.1\,{\rm GeV}^2/c^4$ to $7.0\,{\rm GeV}^2/c^4$. Compared to previous measurements in similar decay modes performed at LHCb, the contamination from $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ decays leaking into the central q^2 region is strongly suppressed by selecting on a novel variable, the constrained q^2. This allows to widen the central q^2 bin up to $7.0\,{\rm GeV}^2/c^4$ instead of $6.0\,{\rm GeV}^2/c^4$, which significantly increases the signal yield in data. The analysis is based on a counting method procedure, which is derived from the symmetry properties of the *CP*-averaged differential decay width of the $B^0 \to K^{*0}\,\ell^+\ell^-$ decay. This approach has the advantage of being very robust and independent of the angular distribution of background contributions. In addition, the measurement involves a novel approach to correct the mismodelling of the kinematics of the decays and the particle identification response of the detector between simulation and data. Machine learning techniques based on the `scikit-learn` [1], `REP` [2,3], and `hep_ml` [4] python packages are used to suppress the background due to random combinations of particles from the same proton-proton collision. The analysis strategy, the description of the samples, the selection requirements, the corrections applied to the simulated candidates, and the efficiency calculation can be found in *Chapter 6*. The mass fits, the cross checks, the systematic uncertainties, and the results are described in *Chapter 7*.

A brief summary of the thesis and a discussion of the future prospects for the upgrade and for the measurement of the angular observables are reported in the *Conclusions*.

CHAPTER 1

Theory

This chapter introduces the theoretical framework of the Standard Model, the theory that attempts to describe Nature at its most fundamental level, and discusses some of its potential New Physics extensions. Focus is mainly directed to rare decays and, in particular, to the $b \to s\ell^+\ell^-$ transition, which is the main subject of the studies described in Chapters 6 and 7.

1.1 The Standard Model

The *Standard Model* (SM) is a *quantum field theory* that describes the fundamental constituents of matter and the interactions among them [5–7]. The model is defined by the symmetries of the Lagrangian and by the representations of the particles under these symmetries. The Lagrangian of the SM is based on the gauge symmetry

$$G_{\rm SM} = SU(3)_{\rm C} \times SU(2)_{\rm L} \times U(1)_{\rm Y},$$

where the $SU(3)_{\rm C}$ and the $SU(2)_{\rm L} \times U(1)_{\rm Y}$ groups describe strong and electroweak interactions, respectively. Strong interactions are the subject of the Quantum ChromoDynamics (QCD), electroweak interactions are the subject of the electroweak theory.

A schematic representation of the SM building blocks is shown in Fig. 1.1. *Quarks* and *leptons* are fermions and appear in three replicas, usually called families or generations: (u, d), (c, s), and (t, b) are the three generations of quarks, while (e, ν_e), (μ, ν_μ), and (τ, ν_τ) are the three generations of leptons. *Gauge bosons* are the mediators of the interactions of quarks and leptons: strong interactions are mediated by eight gluons, electromagnetic interactions are mediated by the photon, and weak interactions are mediated by the Z and $W^\pm$ bosons. The model is completed by the *Higgs boson*, which is a scalar particle related to the spontaneous breaking of the electroweak symmmetry and the origin of particle masses. In a notation that makes the representations and the quantum numbers of the fields manifest, the SM contains three fermion generations, each consisting of five representations

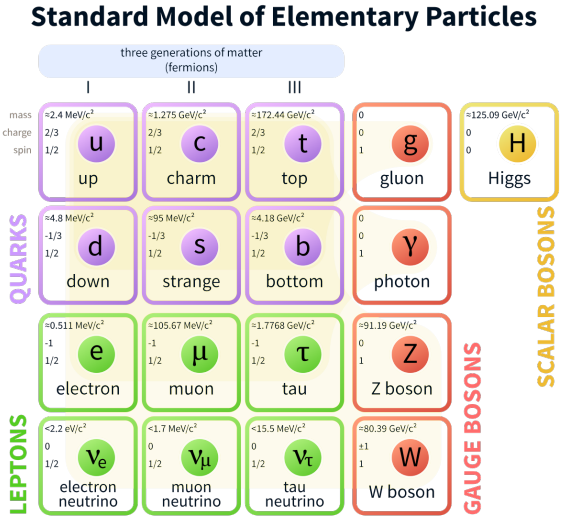


Figure 1.1: Schematic representation of the SM building blocks [8].

of $G_{\rm SM}$,

$$Q^I_{Li}(3,2)_{+1/6}, \quad u^I_{Ri}(3,1)_{+2/3}, \quad d^I_{Ri}(3,1)_{-1/3}, \quad L^I_{Li}(1,2)_{-1/2}, \quad \ell^I_{Ri}(1,1)_{-1},$$

and a single scalar representation, the complex Higgs field,

$$\phi(1,2)_{+1/2}.$$

Left-handed quarks, Q^I_L, are triplets of $SU(3)_{\rm C}$, doublets of $SU(2)_{\rm L}$, and carry hypercharge $Y = +1/6$; right-handed up-type quarks, u^I_R, are triplets of $SU(3)_{\rm C}$, singlets of $SU(2)_{\rm L}$, and carry hypercharge $Y = +2/3$; right-handed down-type quarks, d^I_R, are triplets of $SU(3)_{\rm C}$, singlets of $SU(2)_{\rm L}$, and carry hypercharge $Y = -1/3$. Leptons are singlets of $SU(3)_{\rm C}$ and are classified according to the transformation properties of their fields with respect to $SU(2)_{\rm L}$. Left-handed leptons, L^I_L, are doublets of $SU(2)_{\rm L}$; right-handed leptons, ℓ^I_R, are singlets of $SU(2)_{\rm L}$. The index I denotes interaction eigenstates. The index $i = 1, 2, 3$ identifies the *flavour* (or generation) and corresponds to (u, c, t) for up-type quarks, (d, s, b) for down-type quarks, (e, μ, τ) for charged leptons, and $(\nu_e, \nu_\mu, \nu_\tau)$ for neutral leptons. The complex Higgs field has a vacuum expectation value

$$\langle \phi \rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}$$

that depends on a real free parameter v and generates the spontaneous symmetry breaking of $G_{\rm SM}$ according to the following pattern:

$$G_{\rm SM} \to SU(3)_{\rm C} \times U(1)_{\rm EM}. \tag{1.1}$$

The Z and $W^\pm$ bosons acquire mass through the *Brout-Englert-Higgs mechanism* [9–11], while the photon is massless.

Once the gauge symmetry, the particle content, and the pattern of spontaneous symmetry breaking are defined, the Lagrangian of the SM is derived as a renormalisable Lagrangian satisfying

these requirements. It can be divided in three contributions,

$$\mathcal{L}_{\rm SM} = \mathcal{L}_{\rm kinetic} + \mathcal{L}_{\rm Higgs} + \mathcal{L}_{\rm Yukawa}.$$

The first part, $\mathcal{L}_{\rm kinetic}$, contains the kinetic terms involving covariant derivatives to preserve gauge invariance,

$$\partial^\mu \to D^\mu = \partial^\mu + ig_s G_a^\mu L_a + igW_b^\mu \Sigma_b + ig' B^\mu Y,$$

where

- G_a^μ are the eight gluon fields, W_b^μ are the three weak-interaction fields, and B^μ is the single hypercharge field;
- L_a are the $SU(3)_{\rm C}$ generators (the 3×3 Gell-Mann matrices $\lambda_a/2$ for triplets, 0 for singlets), Σ_b are the $SU(2)_{\rm L}$ generators (the 2×2 Pauli matrices $\sigma_b/2$ for doublets, 0 for singlets), and Y is the $U(1)_{\rm Y}$ charge;
- g_s, g, and g' are the coupling constants of $SU(3)_{\rm C}$, $SU(2)_{\rm L}$, and $U(1)_{\rm Y}$, respectively.

For example, the kinetic term corresponding to left-handed quarks, Q_L^I, is written as

$$\mathcal{L}_{\rm kinetic}(Q_L^I) = i\bar{Q}_{Li}^I \gamma_\mu \left(\partial^\mu + \frac{i}{2} g_s G_a^\mu \lambda_a + \frac{i}{2} g W_b^\mu \tau_b + \frac{i}{6} g' B^\mu \right) Q_{Li}^I, \tag{1.2}$$

while the kinetic term corresponding to right-handed up-type quarks, u_R^I, is written as

$$\mathcal{L}_{\rm kinetic}(u_R^I) = i\bar{u}_{Ri}^I \gamma_\mu \left(\partial^\mu + \frac{i}{2} g_s G_a^\mu \lambda_a + \frac{2}{3} i g' B^\mu \right) u_{Ri}^I, \tag{1.3}$$

since the latter are singlets of $SU(2)_{\rm L}$.

Due to the asymptotic freedom of the strong interactions, quarks can only be observed in color singlet combinations, which are called *hadrons*. These can be pairs of a quark and an antiquark, called *mesons*, triplets of quarks, called *baryons*, or, as observed recently, more exotic states consisting of four or five quarks [12].

The second part, $\mathcal{L}_{\rm Higgs}$, contains the Higgs potential and describes the complex Higgs field self-interactions,

$$\mathcal{L}_{\rm Higgs} = \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2.$$

The parameter μ^2 is required to be negative to satisfy the spontaneous symmetry-breaking pattern of Eq. (1.1), the parameter λ is required to be positive to satisfy vacuum stability.

The third part, $\mathcal{L}_{\rm Yukawa}$, contains the Yukawa terms and describes the interactions of quarks and leptons with the complex Higgs field,

$$-\mathcal{L}_{\rm Yukawa} = Y_{ij}^d \bar{Q}_{Li}^I \phi d_{Rj}^I + Y_{ij}^u \bar{Q}_{Li}^I \tilde{\phi} u_{Rj}^I + Y_{ij}^\ell \bar{L}_{Li}^I \tilde{\phi} \ell_{Rj}^I + \text{h.c.}, \tag{1.4}$$

where $\tilde{\phi} = i\sigma_2 \phi^*$. This part constitutes the flavour sector of the SM, that is, the sector that distinguishes among different generations of quarks and leptons, and is described in detail in the following.

1.1.1 Flavour dynamics

The Yukawa terms of Eq. (1.4) generate quarks and charged leptons masses. This is made explicit by rewriting them in a more convenient form. The complex Higgs field is rewritten as

$$\phi = e^{-i\theta_a(x)\frac{\sigma^a}{2}} \begin{pmatrix} 0 \\ \frac{v+H^0}{\sqrt{2}} \end{pmatrix}, \tag{1.5}$$

where H^0 is a real scalar field, called the "physical Higgs field". The quark and lepton $SU(2)_{\mathrm{L}}$ doublets can be decomposed into their components as

$$Q^I_{Li} = \begin{pmatrix} u^I_{Li} \\ d^I_{Li} \end{pmatrix} \quad L^I_{Li} = \begin{pmatrix} \nu^I_{Li} \\ e^I_{Li} \end{pmatrix}. \tag{1.6}$$

When substituting Eqs. (1.5) and (1.6) into the Yukawa terms of Eq. (1.4), one obtains the interaction terms with the physical Higgs field and the mass terms for quarks and charged leptons,

$$-\mathcal{L}_M = (M_u)_{ij}\bar{u}^I_{Li}u^I_{Rj} + (M_d)_{ij}\bar{d}^I_{Li}d^I_{Rj} + (M_\ell)_{ij}\bar{\ell}^I_{Li}\ell^I_{Rj} + \text{h.c.}, \tag{1.7}$$

where $M_f = \frac{v}{\sqrt{2}}Y^f$ and $f = u, d, \ell$. For quarks, mass matrices are not diagonal in this basis and "mix" fields of different generations. This is related to the difference between interaction eigenstates and mass eigenstates. In terms of mass eigenstates, Eq. (1.7) becomes

$$-\mathcal{L}_M = \bar{u}_L M_u^{\mathrm{diag}} u_R + \bar{d}_L M_d^{\mathrm{diag}} d_R + \bar{\ell}_L M_\ell^{\mathrm{diag}} \ell_R + \text{h.c.},$$

where mass matrices are real and diagonal. Interaction eigenstates and mass eigenstates are related by a transformation from the interaction basis, where interactions are diagonal, to the mass basis, where masses are diagonal. This transformation is realised by two unitary matrices V_{fL} and V_{fR} such that

$$V_{fL}M_f V^\dagger_{fR} = M_f^{\mathrm{diag}},$$

with M_f^{diag} diagonal and real. The mass eigenstates are then identified as

$$d_{Li} = (V_{dL})_{ij}d^I_{Lj}, \quad d_{Ri} = (V_{dR})_{ij}d^I_{Rj}, \tag{1.8}$$

$$u_{Li} = (V_{uL})_{ij}u^I_{Lj}, \quad u_{Ri} = (V_{uR})_{ij}u^I_{Rj}, \tag{1.9}$$

$$\ell_{Li} = (V_{\ell L})_{ij}\ell^I_{Lj}, \quad \ell_{Ri} = (V_{\ell R})_{ij}\ell^I_{Rj}, \quad \nu_{Li} = (V_{\nu L})_{ij}\nu^I_{Lj}, \tag{1.10}$$

where matrices are required to be unitary to preserve the norm of each state.

The charged-current weak interactions, that is, the interactions of the charged $SU(2)_{\mathrm{L}}$ gauge bosons $W^\pm_\mu = (W^1_\mu \mp iW^2_\mu)/\sqrt{2}$) for quarks, which are described by Eqs. (1.2) and (1.3) in the interaction basis, have the following form in the mass basis:

$$-\mathcal{L}_W = \frac{g}{\sqrt{2}}\bar{u}_{Li}\gamma^\mu(V_{uL}V^\dagger_{dL})_{ij}d_{Lj}W^+_\mu + \text{h.c.},$$

where the unitary 3×3 matrix

$$V_{\text{CKM}} = V_{uL} V_{dL}^{\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.11}$$

is the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [13, 14]. This matrix describes the quark transition from one flavour to another through a *Flavour Changing Charged Current* (FCCC) mediated by the W boson. In the SM, FCCCs are the only source of flavour changing interactions. Processes proceeding through *Flavour Changing Neutral Currents* (FCNCs) can only occur at loop level and are hence suppressed in the SM. The CKM matrix of Eq. (1.11) has three real parameters, the mixing angles, and one imaginary parameter, the *CP*-violating phase. The latter is the gateway for *CP* violation in the SM.

Since the SM neutrinos are massless, $V_{\nu L}$ in Eq. (1.10) is arbitrary. Nevertheless, experiments with solar, atmospheric, and reactor neutrinos conducted in recent years established that neutrinos are massive and oscillate among different flavours [15]. This can call for an extension of the SM through either an extension of its Higgs and/or fermion content or an extension of its gauge group (which also requires extended particle content) [16]. One possibility to accommodate neutrino masses and oscillations is described in the next section.

1.1.2 Lepton flavour conservation and universality

Three lepton-flavour quantum numbers corresponding to the three generations of leptons can be defined. These quantum numbers, referred to as Q_e, Q_μ, and Q_τ, can have values of $+1$, -1, or 0, depending on the flavour and charge of the particle fields. For example, the electron quantum number Q_e has a value of $+1$ for the e^- field, a value of -1 for the e^+ field, and a value of 0 for all remaining fields. Similar considerations apply to Q_μ and Q_τ. Since the Lagrangian of the SM is invariant under global rotations of the lepton fields of the type $U(1)_e \times U(1)_\mu \times U(1)_\tau$, individual lepton-flavour quantum numbers are conserved. This conservation law is known as *Lepton Flavour Conservation* (LFC). As a consequence, also the total lepton-flavour quantum number $Q_L = Q_e + Q_\mu + Q_\tau$, corresponding to global rotations of the type $U(1)_L$, is conserved. However, neutrino oscillations are a clear experimental evidence of *Neutral Lepton Flavour Violation* (NLFV) [15]. This can be accommodated in the SM by requiring at least two neutrinos to be massive and to have distinct mass values and by introducing a 3×3 unitary neutrino-mixing matrix, the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [17, 18], similar to the quark-mixing matrix described in the previous section. This is achieved by adding three right-handed neutrino fields N_{Ri}^I with the following Lagrangian:

$$\Delta\mathcal{L} = i\bar{N}_{Ri}^I \gamma_\mu D^\mu N_{Ri}^I - Y_{ij}^\nu \bar{L}_{Li}^I \tilde{\phi} N_{Rj}^I + \text{h.c.},$$

so that quarks and leptons have an analogous flavour structure. In addition to NLFV, many extensions of the SM foresee the possibility to have *Charged Lepton Flavour Violation* (CLFV) [15], that is, lepton-flavour quantum number violation in the transition between two charged leptons. The main experimental efforts in this direction are the searches for the forbidden decays $\mu \to e\gamma$ [19], $\mu \to eee$ [20], and $\mu N \to eN$ [21–23].

Apart from the Yukawa terms, the Lagrangian of the SM is lepton-flavour universal. In particular,

Lepton Flavour Universality (LFU) states that the electroweak couplings of the three generations of leptons are the same (*i.e.*, electrons couple to photons, Z and $W^{\pm}$ bosons the same way muons and taus do). However, recent measurements have shown hints of deviations from the SM predictions that could point to models in which LFU is violated. The experimental status of these searches, which is central for the topics described in this thesis, is described in more detail in Section 1.6.

1.2 Beyond the Standard Model

Although the SM provides a satisfactory explanation of interactions among elementary particles, some open questions still remain. The SM does not provide an explanation on why there are three fermion generations, nor why their masses have values spreading over more than twelve orders of magnitude. Moreover, it does not describe the dark-matter content of the Universe and does not include gravity. These and other unanswered questions suggest that a more fundamental theory, generically referred to as *New Physics* (NP), is likely to exist. In this scenario, the SM is interpreted as an effective theory corresponding to the low-energy approximation of such a more fundamental theory.

Several extensions of the SM have been proposed, differing in both their particle content and interactions: they are all designed to reduce to the SM at relatively low energies, but offer solutions to its shortcomings and predict new physical phenomena at higher energies, where the effects of the added degrees of freedom would become detectable. Two NP models that are especially relevant for the measurements described in this thesis are described in Section 1.5.

Experimentally, the efforts toward identifying and characterising the possible extensions of the SM can be broadly classified in two complementary approaches. *Direct searches* aim at directly producing and observing new particles. This approach has the advantage of allowing a direct observation of new particles, but its reach is limited by the highest energies that can be achieved. *Indirect searches* play an important and complementary role, exploiting the possibility that contributions from NP may arise in virtual loops of amplitudes involving new particles or interactions. This may induce differences between experimentally observed results and corresponding SM predictions, possibly evidencing the presence of NP. Thus, indirect searches are sensitive also to NP at energy scales far beyond those accessible to present particle accelerators. In the framework of indirect searches, independent measurements of a large number of processes can be performed and then compared with one another seeking any inconsistency that could hint at NP. The study of rare decays of hadrons containing a b quark, in particular, is one of the most promising approaches to search for NP effects and is described in more detail in the next section.

1.3 Rare decays of beauty hadrons

Processes in which a quark of a given generation transforms into one of a different generation can happen through charged or neutral currents. In the SM, the exchange of a W boson represents a FCCC, which can therefore occur at tree level. Conversely, processes proceeding through FCNCs can only occur at loop level and are predicted to be rare. However, NP contributions can result in sizeable deviations from the SM predictions, which makes FCNC processes a particularly powerful tool to probe flavour structures beyond the SM. The main issue is to connect quantitatively and

precisely the predictions from a quark-based theory with hadron-based measurements. Theoretically, the main source of uncertainty originates from low-energy strong interactions, which modify the purely electroweak amplitudes in ways that are challenging to calculate. Depending on the kinematic region of the process, however, several approximations can be used to quantity these QCD effects. Among FCNC processes, rare decays of beauty hadrons are particularly interesting since the large mass of the b quark with respect to the energy scale of QCD ($m_b \gg \Lambda_{\rm QCD}$) allows to separate strong and electroweak contributions and predict decay rates and other properties with small theoretical uncertainty. Typical branching ratios for these decays are below 10^{-6}, which explains why they are still partially unexplored. The rest of the chapter will focus on one specific type of rare decays of beauty hadrons, namely those proceeding through a $b \to s\ell^+\ell^-$ transition.

1.4 Effective field theory

From the phenomenological point of view, experimental results can be interpreted by following a model-independent approach in the framework of a low-energy *effective field theory*. The latter is obtained by integrating out the degrees of freedom of the particles that are much heavier than the b quark, namely the t quark and the $W^\pm$ and Z bosons. In a similar fashion to Fermi's theory of weak decays, the Hamiltonian of the full theory describing the $b \to s$ transition is hence replaced by an effective Hamiltonian [24] of the form

$$\mathcal{H}_{\rm eff} = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_i \left(\mathcal{C}_i(\mu_s)\mathcal{O}_i(\mu_s) + \mathcal{C}'_i(\mu_s)\mathcal{O}'_i(\mu_s)\right),$$

where G_F is the Fermi constant, V_{tb} and V_{ts} are the elements of the CKM matrix, $\mathcal{C}_i$ and $\mathcal{C}'_i$ are the Wilson coefficients describing the short-range contribution, and $\mathcal{O}_i$ and $\mathcal{O}'_i$ are the local operators encoding the long-range contribution. Although particles heavier than the b quark are removed from the theory as dynamical degrees of freedom, the effect of their existence is implicitly taken into account in the Wilson coefficients. Doubly Cabibbo-suppressed contributions to the effective Hamiltonian of the type $V_{ub}V_{us}^*$ are neglected in the calculation.

The Wilson coefficients and the local operators are evaluated at the renormalisation scale μ_s. All particles contributing to the $b \to s$ transition and having a mass larger than the renormalisation scale affect the value of at least one Wilson coefficient. The values of the Wilson coefficients at the electroweak scale are obtained by matching the decay amplitudes of the full theory with those of the effective theory. Renormalisation group equations are then used to evolve the Wilson coefficients from the electroweak to the renormalisation scale [25, 26]. Assuming SM dynamics, the non-vanishing Wilson coefficients have the following values at $\mu_s = m_b$, as obtained from a Next-to-Next-to-Leading-Logarithm (NNLL) calculation [27, 28]:

$$\mathcal{C}_7^{\rm SM} = -0.3, \quad \mathcal{C}_9^{\rm SM} = +4.2, \quad \mathcal{C}_{10}^{\rm SM} = -4.2.$$

Contributions of NP appear in the Wilson coefficients as additive factors:

$$\mathcal{C}_i = \mathcal{C}_i^{\rm SM} + \mathcal{C}_i^{\rm NP}.$$

The amplitude of a given process can be computed as the expectation value of the effective

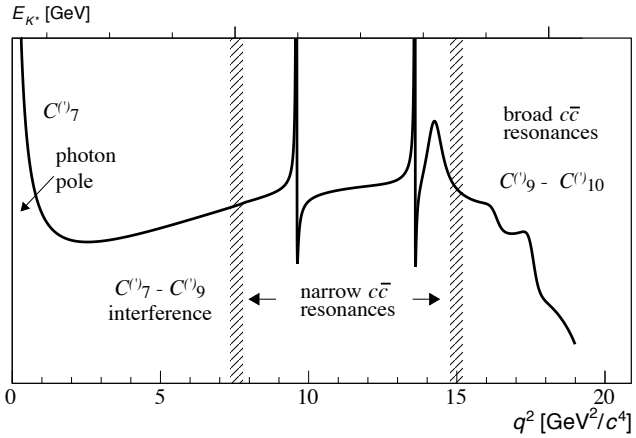


Figure 1.2: Qualitative dependence of the $B^0 \to K^{*0}\ell^+\ell^-$ differential decay width on q^2. The two narrow peaks correspond to the J/ψ and $\psi(2S)$ resonances, while the wide peaks on the right side correspond to the broad charmonia resonances. Figure modified from Ref. [29].

Hamiltonian between the initial and final states:

$$A(I \to F) = \langle F|\mathcal{H}_{\text{eff}}|I\rangle = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_i \left(\mathcal{C}_i \langle F|\mathcal{O}_i|I\rangle + \mathcal{C}'_i \langle F|\mathcal{O}'_i|I\rangle\right),$$

where I and F denote respectively the initial and final states. The hadronic matrix elements $\langle F|\mathcal{O}_i|I\rangle$ and $\langle F|\mathcal{O}'_i|I\rangle$ contain information about the *form factors*, which represent the dominant source of uncertainty in the predictions.

The local operators differ in their Lorentz structure. The following dimension-six local operators are relevant to describe the $b \to s\ell^+\ell^-$ transition:

$$\begin{array}{rclcrcl} \mathcal{O}_7 &=& \frac{m_b}{e}\left[\bar{s}\sigma^{\mu\nu}P_R b\right] F_{\mu\nu}\,, & & \mathcal{O}'_7 &=& \frac{m_b}{e}\left[\bar{s}\sigma^{\mu\nu}P_L b\right] F_{\mu\nu}\,, \\ \mathcal{O}_9 &=& \left[\bar{s}\gamma_\mu P_L b\right]\left[\bar{\ell}\gamma^\mu \ell\right]\,, & & \mathcal{O}'_9 &=& \left[\bar{s}\gamma_\mu P_R b\right]\left[\bar{\ell}\gamma^\mu \ell\right]\,, \\ \mathcal{O}_{10} &=& \left[\bar{s}\gamma_\mu P_L b\right]\left[\bar{\ell}\gamma^\mu\gamma_5 \ell\right]\,, & & \mathcal{O}'_{10} &=& \left[\bar{s}\gamma_\mu P_R b\right]\left[\bar{\ell}\gamma^\mu\gamma_5 \ell\right]\,. \end{array} \tag{1.12}$$

In the formulas above, m_b is the mass of the b quark, e is the charge of the electron, s and b are the fields associated to the s and b quarks, respectively, ℓ is the field associated to the lepton ℓ, $F_{\mu\nu}$ is the electromagnetic tensor, and $P_{L/R} = (1 \mp \gamma_5)/2$ are the left/right handed chiral projectors. Chirality-flipped operators $\mathcal{O}'$ are obtained by interchanging P_L and P_R in the quark currents. The $\mathcal{O}_7$ operator, which is referred to as *electromagnetic operator*, is the dominant contribution to the radiative $b \to s\gamma$ transition. The $\mathcal{O}_9$ and $\mathcal{O}_{10}$ operators, which are called *semileptonic operators*, receive contributions from penguin and box diagrams mediated by Z and $W^\pm$ bosons. In addition to the operators in Eq. (1.12), the following *scalar* and *pseudoscalar operators* can be defined:

$$\begin{array}{rclcrcl} \mathcal{O}_{\text{S}} &=& \left[\bar{s}P_R b\right]\left[\bar{\ell}\ell\right]\,, & & \mathcal{O}'_{\text{S}} &=& \left[\bar{s}P_L b\right]\left[\bar{\ell}\ell\right]\,, \\ \mathcal{O}_{\text{P}} &=& \left[\bar{s}P_R b\right]\left[\bar{\ell}\gamma_5 \ell\right]\,, & & \mathcal{O}'_{\text{P}} &=& \left[\bar{s}P_L b\right]\left[\bar{\ell}\gamma_5 \ell\right]\,. \end{array} \tag{1.13}$$

These are highly suppressed in the SM due to the small mass of the leptons [30]. Contributions from

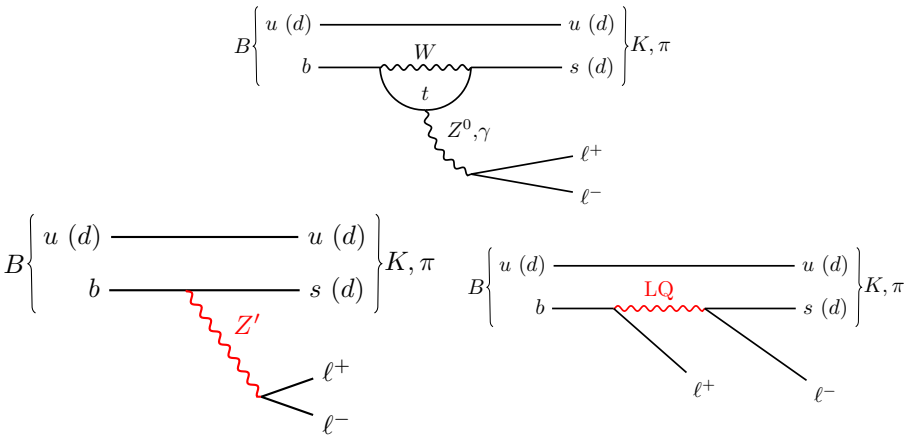


Figure 1.3: Feynman diagrams contributing to the $b \to s(d)\ell^+\ell^-$ transition at the lowest order of perturbation theory in the SM (top) and in NP models violating LFU, such as models involving Z' bosons (bottom left) or leptoquarks (bottom right).

NP models can also originate from the following *tensor operators*:

$$\mathcal{O}_{\rm T} = [\bar{s}\sigma_{\mu\nu}b]\left[\bar{\ell}\sigma^{\mu\nu}\ell\right], \quad \mathcal{O}_{\rm T5} = [\bar{s}\sigma_{\mu\nu}b]\left[\bar{\ell}\sigma^{\mu\nu}\gamma_5\ell\right], \tag{1.14}$$

which are all negligibly small in the SM [30]. Due to the left-handed nature of the weak interactions, the $\mathcal{C}'_7$ right-handed Wilson coefficient is suppressed in the SM, while the $\mathcal{C}'_9$ and $\mathcal{C}'_{10}$ right-handed Wilson coefficients are zero. Therefore, $\mathcal{C}'_7$, $\mathcal{C}'_9$, and $\mathcal{C}'_{10}$ can receive non-negligible contributions from NP models involving a different helicity structure [31, 32].

Radiative decays of the type $B^0 \to K^{*0}\gamma$ receive contributions from $\mathcal{C}_7^{(\prime)}$ only, leptonic decays of the type $B^0 \to \ell^+\ell^-$ receive contributions from $\mathcal{C}_{10}^{(\prime)}$, while semileptonic decays of the type $B^0 \to K^{*0}\ell^+\ell^-$ receive contributions from $\mathcal{C}_7^{(\prime)}$, $\mathcal{C}_9^{(\prime)}$, and $\mathcal{C}_{10}^{(\prime)}$. Different measurements are hence sensitive to different combinations of local operators and can be compared with one another to provide consistency checks and classify NP contributions that are compatible with the observations. The distribution in Fig. 1.2 describes the qualitative dependence of the $B^0 \to K^{*0}\ell^+\ell^-$ differential decay width on q^2, where q^2 is the dilepton invariant mass squared. Two differerent kinematic regimes are visible: the low q^2 region, where the emitted hadron is energetic, and the high q^2 region, where the emitted hadron has low recoil energy. For low values of q^2, up to 1 GeV$^2/c^4$, the main contribution originates from the $\mathcal{O}_7$ and $\mathcal{O}'_7$ operators, which are related to the $b \to s\gamma$ transition. For values of q^2 between 1 and 6 GeV$^2/c^4$ the interference among $\mathcal{O}_7^{(\prime)}$, $\mathcal{O}_9^{(\prime)}$, and $\mathcal{O}_{10}^{(\prime)}$ plays a major role.

1.5 The $b \to s\ell^+\ell^-$ transition

As mentioned in the previous section, the decays proceeding through a $b \to s\ell^+\ell^-$ transition are suppressed in the SM due to the absence of FCNCs at tree level. The contribution at the lowest order of perturbation theory originates from one-loop processes consisting of box and penguin diagrams, of which one example is shown in Fig. 1.3 (top). Proceeding through loop processes, these decays are highly sensitive to the possible presence of virtual particles and interactions predicted in many extensions of the SM, like those in which leptoquarks [15] and Z' bosons [15] mediate the transition from the b quark to the s quark, as shown in Fig. 1.3 (bottom). A Z' boson is a neutral particle

of spin 1 that originates from an extra $U(1)$ gauge symmetry, for example $U(1)_{\mathrm{B-L}}$, where B and L are the baryon and lepton quantum numbers. If the coupling of the Z' boson to the leptons is not universal for the three generations, LFU breaking occurs. A leptoquark is a boson (typically of spin 0, although NP models with leptoquarks of spin 1 exist) that carries both quark and lepton quantum numbers. A tree level exchange of a leptoquark may occur in processes such as $b \to s\ell^+\ell^-$ and can therefore enhance the decay rates of such processes with respect to their SM predictions. Leptoquarks would also provide a natural source of LFU violation.

The extensions of the SM that foresee a LFU violation can be probed by comparing decays with muons and electrons in the final state.

This comparison can be achieved by measuring ratios of the type

$$R_H = \frac{\int \frac{d\Gamma(B\to H\mu^+\mu^-)}{dq^2} dq^2}{\int \frac{d\Gamma(B\to He^+e^-)}{dq^2} dq^2},$$

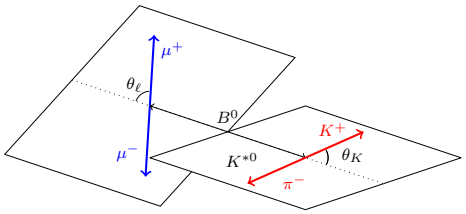
where Γ is the decay width, q^2 is the dilepton invariant mass squared, and H is a hadron, such as a K or K^* meson. The advantage of this approach is that very precise SM predictions are available. Two of these measurements, referred to as R_K and R_{K^*} in literature, were performed recently at LHCb [33,34] and are briefly described in Section 1.6.

A different approach, complementary to the measurement of ratios like R_K and R_{K^*}, consists in studying the angular distribution of the particles in the final state for decays to leptons of different generations. Due to the presence of a vector meson in the final state, $B^0 \to K^{*0}\ \ell^+\ell^-$ decays, with the K^{*0} reconstructed in the $K^+\pi^-$ final state, offer a particularly rich phenomenology and are the main topic of the rest of the section. The angular distribution of the particles in the final state can be expressed in terms of q^2 and three *decay angles* $\vec{\Omega} = (\cos\theta_\ell, \cos\theta_K, \phi)$ [35]. The definition of the decay angles is shown in Fig. 1.4. The angle θ_ℓ is the angle between the direction of the ℓ^+ (ℓ^-) and the direction opposite to that of the B^0 ($\overline{B}^0$) in the dilepton reference frame. The angle θ_K is the angle between the direction of the K^+ (K^-) and the direction of the B^0 ($\overline{B}^0$) in the K^{*0} ($\overline{K}^{*0}$) reference frame. The angle ϕ is the angle between the plane of the dilepton pair and the plane of the K and π originated from the K^{*0} ($\overline{K}^{*0}$) decay in the B^0 ($\overline{B}^0$) reference frame.

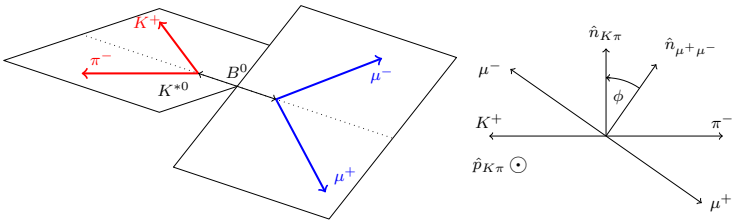
The differential decay widths of the $\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays in terms of q^2 and $\vec{\Omega}$ are given by

$$\begin{aligned} \frac{\mathrm{d}^4\Gamma[\overline{B}^0 \to \overline{K}^{*0}\ell^+\ell^-]}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}} &= \frac{9}{32\pi}\sum_i I_i(q^2) f_i(\vec{\Omega}) \quad \text{and} \\ \frac{\mathrm{d}^4\bar{\Gamma}[B^0 \to K^{*0}\ell^+\ell^-]}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}} &= \frac{9}{32\pi}\sum_i \bar{I}_i(q^2) f_i(\vec{\Omega}) \ , \end{aligned} \tag{1.15}$$

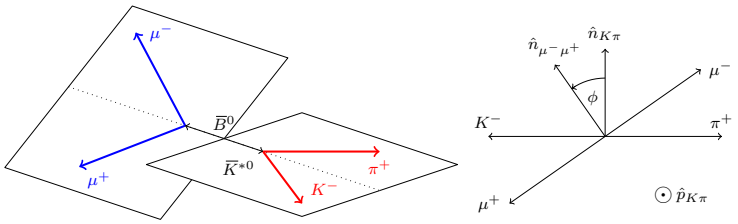
where Γ ($\bar{\Gamma}$) refers to decays involving a b ($\bar{b}$) quark and hence a $\overline{B}^0$ (B^0) meson, the terms $f_i(\vec{\Omega})$ are combinations of spherical harmonics and the terms I_i ($\bar{I}_i$) are q^2-dependent *angular observables*. In the limit of massless leptons, corresponding to $m^2_{\ell^+\ell^-} \gg 4\,\mathrm{m}^2_\ell$, the latter can be expressed as bilinear combinations of six complex decay amplitudes, $\mathcal{A}^{\mathrm{L,R}}_{0,\parallel,\perp}$, which correspond to the three different transversity states of the K^{*0} meson (one longitudinal polarisation and two transverse polarisations) and two different chiralities (left-handed or right-handed) of the dilepton system. It should be noted here that the timelike amplitude $\mathcal{A}_t$ is not considered since it vanishes in the limit of massless leptons.



(a) θ_K and θ_ℓ definitions for the B^0 decay



(b) ϕ definition for the B^0 decay



(c) ϕ definition for the $\overline{B}^0$ decay

Figure 1.4: Definition of the decay angles θ_ℓ, θ_K, and ϕ describing the angular distribution of the decay products of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decay.

An additional suffix s or c is usually added to some of the angular observables to indicate that they have a dependence on $\sin^2\theta_K$ or $\cos^2\theta_K$. The list of the $I_i(q^2)$ and $f_i(\vec{\Omega})$ terms that remain in the limit of massless leptons is reported in Table 1.1.

The differential decay widths of Eq. (1.15) can be combined to form *CP* averages S_i and *CP* asymmetries A_i observables, according to the following:

$$\begin{aligned} S_i &= \left(I_i + \bar{I}_i\right) \Big/ \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}\right) \text{ and} \\ A_i &= \left(I_i - \bar{I}_i\right) \Big/ \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}\right). \end{aligned} \tag{1.16}$$

Two S_i observables, S_{1c} and S_{6s}, have a simple physical interpretation. The S_{1c} observable corresponds to the *fraction of longitudinal polarisation* of the K^{*0} meson and is therefore more commonly referred to as F_L, with

$$F_\mathrm{L} = S_{1c} = \frac{|\mathcal{A}_0^\mathrm{L}|^2 + |\mathcal{A}_0^\mathrm{R}|^2}{|\mathcal{A}_0^\mathrm{L}|^2 + |\mathcal{A}_0^\mathrm{R}|^2 + |\mathcal{A}_\parallel^\mathrm{L}|^2 + |\mathcal{A}_\parallel^\mathrm{R}|^2 + |\mathcal{A}_\perp^\mathrm{L}|^2 + |\mathcal{A}_\perp^\mathrm{R}|^2}\,.$$

The S_{6s} observable is related to the *forward-backward asymmetry* of the dilepton system and is conventionally replaced by $A_{\rm FB}$, with $A_{\rm FB} = \frac{3}{4}S_{6s}$.

After substituting $F_{\rm L}$ and $A_{\rm FB}$ in Eq. (1.16), the *CP*-averaged differential decay width of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decay can be written as

$$\begin{aligned}\frac{1}{d(\Gamma+\bar{\Gamma})/dq^2}\frac{d^4(\Gamma+\bar{\Gamma})}{d\vec{\Omega}dq^2} &= \frac{9}{32\pi}[\tfrac{3}{4}(1-F_L)\sin^2\theta_k + F_L\cos^2\theta_k \\ &+\tfrac{1}{4}(1-F_L)\sin^2\theta_k\cos 2\theta_\ell - F_L\cos^2\theta_k\cos 2\theta_\ell \\ &+S_3\sin^2\theta_k\sin^2\theta_\ell\cos 2\phi + S_4\sin 2\theta_k\sin 2\theta_\ell\cos\phi \\ &+S_5\sin 2\theta_k\sin\theta_\ell\cos\phi + \tfrac{4}{3}A_{FB}\sin^2\theta_k\cos\theta_\ell \\ &+S_7\sin 2\theta_k\sin\theta_\ell\sin\phi + S_8\sin 2\theta_k\sin 2\theta_\ell\sin\phi \\ &+S_9\sin^2\theta_k\sin^2\theta_\ell\sin 2\phi],\end{aligned} \tag{1.17}$$

where the highlighted parameters are the angular observables measured in Chapters 6 and 7. The angular observables are combinations of the K^{*0} decay amplitudes and are sensitive to the Wilson coefficients $\mathcal{C}_7^{(\prime)}$, $\mathcal{C}_9^{(\prime)}$, and $\mathcal{C}_{10}^{(\prime)}$ described in Section 1.4. These angular observables can potentially have large theoretical uncertainties due to the limited knowledge of the hadronic form factors. However, optimised variables with reduced theoretical uncertainties can be defined. These are combinations of the $F_{\rm L}$ and $S_3 - S_9$ observables, for example

$$P_5' = \frac{S_5}{\sqrt{F_L(1-F_L)}},$$

which gives the name to the so-called P_5' *anomaly* discussed in the next section. Additional optimised variables are the *transverse asymmetry* $A_{\rm T}^{(2)}$ [36], where $A_{\rm T}^{(2)} = 2S_3/(1-F_{\rm L})$, and the $P_i^{(\prime)}$ series of angular observables described in Ref. [37]. Experimentally, the angular observables can be computed by performing a maximum likelihood fit of the angular distribution of the final state particles or by

Table 1.1: Non-negligible $I_i(q^2)$ and $f_i(\vec{\Omega})$ in the massless limit. The terms arising from the S-wave contribution to the $K^+\pi^-\ell^+\ell^-$ final state are not shown.

i	I_i	f_i
$1s$	$\frac{3}{4}\left[\lvert\mathcal{A}_\parallel^{\rm L}\rvert^2 + \lvert\mathcal{A}_\perp^{\rm L}\rvert^2 + \lvert\mathcal{A}_\parallel^{\rm R}\rvert^2 + \lvert\mathcal{A}_\perp^{\rm R}\rvert^2\right]$	$\sin^2\theta_K$
$1c$	$\lvert\mathcal{A}_0^{\rm L}\rvert^2 + \lvert\mathcal{A}_0^{\rm R}\rvert^2$	$\cos^2\theta_K$
$2s$	$\frac{1}{4}\left[\lvert\mathcal{A}_\parallel^{\rm L}\rvert^2 + \lvert\mathcal{A}_\perp^{\rm L}\rvert^2 + \lvert\mathcal{A}_\parallel^{\rm R}\rvert^2 + \lvert\mathcal{A}_\perp^{\rm R}\rvert^2\right]$	$\sin^2\theta_K\cos 2\theta_l$
$2c$	$-\lvert\mathcal{A}_0^{\rm L}\rvert^2 - \lvert\mathcal{A}_0^{\rm R}\rvert^2$	$\cos^2\theta_K\cos 2\theta_l$
3	$\frac{1}{2}\left[\lvert\mathcal{A}_\perp^{\rm L}\rvert^2 - \lvert\mathcal{A}_\parallel^{\rm L}\rvert^2 + \lvert\mathcal{A}_\perp^{\rm R}\rvert^2 - \lvert\mathcal{A}_\parallel^{\rm R}\rvert^2\right]$	$\sin^2\theta_K\sin^2\theta_l\cos 2\phi$
4	$\sqrt{\frac{1}{2}}{\rm Re}(\mathcal{A}_0^{\rm L}\mathcal{A}_\parallel^{\rm L*} + \mathcal{A}_0^{\rm R}\mathcal{A}_\parallel^{\rm R*})$	$\sin 2\theta_K\sin 2\theta_l\cos\phi$
5	$\sqrt{2}{\rm Re}(\mathcal{A}_0^{\rm L}\mathcal{A}_\perp^{\rm L*} - \mathcal{A}_0^{\rm R}\mathcal{A}_\perp^{\rm R*})$	$\sin 2\theta_K\sin\theta_l\cos\phi$
$6s$	$2{\rm Re}(\mathcal{A}_\parallel^{\rm L}\mathcal{A}_\perp^{\rm L*} - \mathcal{A}_\parallel^{\rm R}\mathcal{A}_\perp^{\rm R*})$	$\sin^2\theta_K\cos\theta_l$
7	$\sqrt{2}{\rm Im}(\mathcal{A}_0^{\rm L}\mathcal{A}_\parallel^{\rm L*} - \mathcal{A}_0^{\rm R}\mathcal{A}_\parallel^{\rm R*})$	$\sin 2\theta_K\sin\theta_l\sin\phi$
8	$\sqrt{\frac{1}{2}}{\rm Im}(\mathcal{A}_0^{\rm L}\mathcal{A}_\perp^{\rm L*} + \mathcal{A}_0^{\rm R}\mathcal{A}_\perp^{\rm R*})$	$\sin 2\theta_K\sin 2\theta_l\sin\phi$
9	${\rm Im}(\mathcal{A}_\parallel^{\rm L*}\mathcal{A}_\perp^{\rm L} + \mathcal{A}_\parallel^{\rm R*}\mathcal{A}_\perp^{\rm R})$	$\sin^2\theta_K\sin^2\theta_l\sin 2\phi$

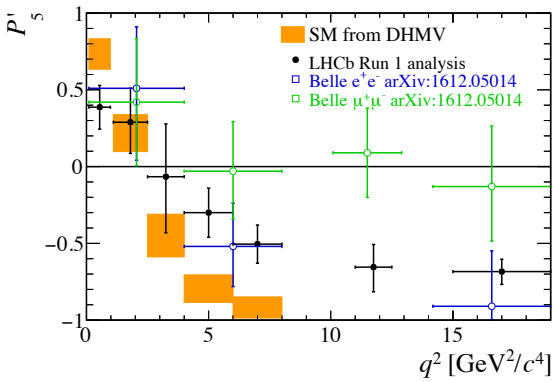


Figure 1.5: Values of P'_5 measured by LHCb (black points) compared to SM predictions (orange bands). Overlaid, values of P'_5 measured by Belle in the individual measurements with only muons (green points) or electrons (blue points) in the final state [38].

applying the method of moments. Both approaches are described in detail in Chapter 6.

In addition to the resonant P-wave contribution from K^{*0} to the $K^+\pi^-\ell^+\ell^-$ final state, the $K^+\pi^-$ system can also be in an S-wave configuration. The addition of the S-wave component introduces two new complex amplitudes, $\mathcal{A}_S^{L,R}$, and results in additional angular terms in the differential decay width. The addition of the S-wave component causes a change of nearly 5% in the differential decay width. This contribution can be taken into account when computing the angular observables of muon decay modes, while electron decay modes do not have at the moment sufficient statistics to include this contribution in the measurements.

According to LFU, no difference is expected between angular observables of muon and electron decay modes. A deviation from the SM predictions would hence be a clear hint of NP effects related to LFU violation.

1.6 Experimental results

As described above, there are two main approaches to search for NP effects and, in particular, for LFU violation in the $b\to s\ell^+\ell^-$ transition. The first approach consists in measuring LFU ratios like R_K and R_{K^*}. The second investigates the angular dependence of the differential decay width of $b\to s\ell^+\ell^-$ decays to muons and electrons. In recent years, several measurements based on both approaches have shown hints of tension with respect to the SM predictions.

One of the most interesting measurements is the P'_5 anomaly observed by LHCb and then confirmed by Belle. The measurement of the angular distribution of the decay products of the $B^0\to K^{*0}\mu^+\mu^-$ decay was performed by LHCb in 8 bins of q^2 from $0.1\,\text{GeV}^2/c^4$ to $19.0\,\text{GeV}^2/c^4$. A deviation was observed in the angular distribution and, in particular, in the P'_5 observable, which is particularly sensitive to the $\mathcal{C}_9$ Wilson coefficient [35]. Incidentally, a compatible deviation was observed by Belle, which performed a measurement in 4 bins of q^2 in the same kinematic region of LHCb. The plot in Fig. 1.5 shows the results obtained by LHCb and Belle, with the latter given separately for decay modes with muons and, respectively, electrons in the final state. The excess is observed in two bins of q^2, respectively from $4.0\,\text{GeV}^2/c^4$ to $6.0\,\text{GeV}^2/c^4$ (at a level of 2.8 standard deviations) and from $6.0\,\text{GeV}^2/c^4$ to $8.0\,\text{GeV}^2/c^4$ (at a level of 3.0 standard deviations).

A similar measurement for the $B^0\to K^{*0}e^+e^-$ decay was performed by the LHCb experiment

in the low q^2 region, from $0.0020 \pm 0.0008\,\text{GeV}^2/c^4$ to $1.120 \pm 0.060\ \text{GeV}^2/c^4$ [39]. This region is dominated by the $b \to s\gamma$ transition and is hence particularly suitable to measure the photon polarisation and the C_7 and C_7' Wilson coefficients. The angular observables were obtained by fitting the B^0 invariant mass distribution and the three decay angles. The results are dominated by statistical uncertainties and are compatible with the SM predictions [40,41].

Hints of deviations from the SM predictions were observed also in the measurements of the LFU ratios R_K and R_{K^*} performed at LHCb. Both were previously measured by BaBar and Belle, which found them compatible with the SM predictions, although within a large statistical uncertainty [42,43]. The first measurement, R_K, is the ratio of the branching fractions of the $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ decays. This is predicted to be unity within an uncertainty of $\mathcal{O}(10^{-2})$ in the SM [30,44]. However, discrepancies from this prediction are expected in several NP scenarios, involving new scalar or pseudoscalar interactions [30] or the existence of a Z' boson coupling differently to electrons and muons [45–47]. The value of R_K was measured by LHCb in the q^2 region between $1\,\text{GeV}^2/c^4$ and $6\ \text{GeV}^2/c^4$ [33]. This region is both experimentally and theoretically favoured, since it allows to veto decays occurring through a J/ψ resonance and to profit from accurate predictions. This resulted in

$$R_K = 0.745^{+0.090}_{-0.074} \pm 0.036,$$

where the first uncertainty is statistical and the second systematic. This result, which corresponds to the most precise measurement to date, is compatible with the SM prediction at 2.6 standard deviations. The second measurement, R_{K^*}, is analogous to the first, but is obtained by studying the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays. The main difference is the presence of a vector meson in the final state. As for R_K, the SM prediction is known at the percent level and differs from unity mainly because of phase-space effects. The R_{K^*} measurement can probe models with leptoquarks or Z' bosons. At LHCb, R_{K^*} was measured in two different regions of q^2, a low q^2 bin between $0.045\,\text{GeV}^2/c^4$ and $1.1\,\text{GeV}^2/c^4$, where the measurement is mostly sensitive to $\mathcal{C}_7$ and $\mathcal{C}_7'$, and a central q^2 bin between $1.1\,\text{GeV}^2/c^4$ and $6\,\text{GeV}^2/c^4$, where a large sensitivity is expected to $\mathcal{C}_9^{(\prime)}$ and $\mathcal{C}_{10}^{(\prime)}$. The measured values are

$$R_{K^{*0}} = \begin{cases} 0.66\ ^{+\ 0.11}_{-\ 0.07}\,(\text{stat}) \pm 0.03\,(\text{syst}) & \text{for } 0.045 < q^2 < 1.1\ \ \text{GeV}^2/c^4 \\ 0.69\ ^{+\ 0.11}_{-\ 0.07}\,(\text{stat}) \pm 0.05\,(\text{syst}) & \text{for } 1.1\ \ \ \ < q^2 < 6.0\ \ \text{GeV}^2/c^4 \end{cases}$$

and correspond to the most precise measurement of R_{K^*} to date. The results are compatible with the SM predictions at $2.1 - 2.3$ and $2.4 - 2.5$ standard deviations in the low and in the central q^2 bin, respectively.

In addition to these measurements, discrepancies with respect to the SM predictions were observed in several branching ratios of $b \to s\mu^+\mu^-$ transitions [48].

The $b \to s\ell^+\ell^-$ measurements described in this section show an interesting and coherent pattern of deviations with respect to the SM predictions. A brief discussion about how to interpret these results is given below.

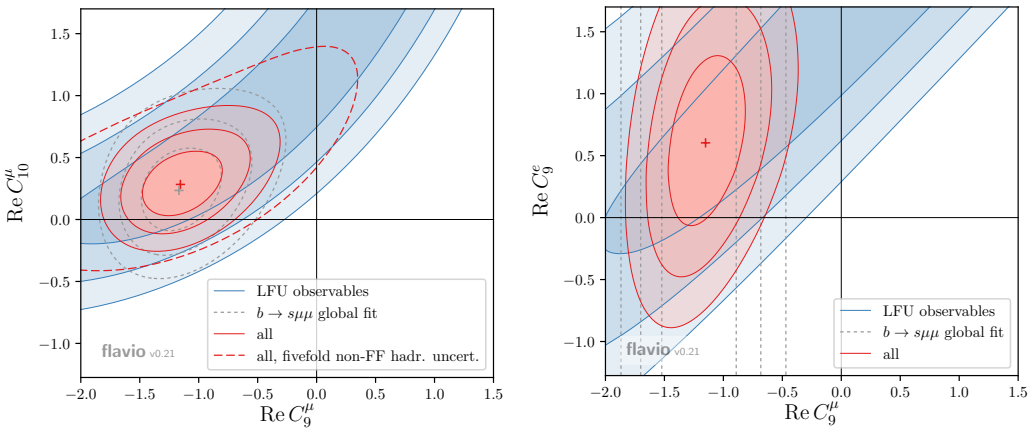


Figure 1.6: Allowed regions in the $(\mathcal{C}_9^{\mu}, \mathcal{C}_{10}^{\mu})$ (left) and $(\mathcal{C}_9^{\mu}, \mathcal{C}_9^{e})$ (right) planes of Wilson coefficients, assuming the remaining Wilson coefficients to be SM-like [47, 49].

1.7 Phenomenological interpretation

The anomaly observed in the angular distribution of the decay products of the $B^0 \to K^{*0}\mu^+\mu^-$ decay might be the result of the underestimation of the theoretical uncertainty or the poor theoretical knowledge of the contribution of the $c\bar{c}$ resonances to the differential decay width of the $B^0 \to K^{*0}\mu^+\mu^-$ decay. While this might explain the excess observed in the P_5' observable of the $B^0 \to K^{*0}\mu^+\mu^-$ decay, a different explanation is required to explain the deviations observed in the R_K and R_{K^*} measurements.

If, on the other hand, one assumes that the deviations are the result of NP, the effective field theory described in Section 1.4 can help in clarifying what kind of NP models might be consistent with the experimental observations.

Several global fits were performed, using as experimental input the R_K and R_{K^*} measurements, as well as the $b \to s\mu^+\mu^-$ measurements. In the global fits, NP contributions that might explain the experimental anomalies were searched for among the dimension-six operators of Eqs. (1.12), (1.13), and (1.14), that is, the vector and axial-vector operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ (where vector and axial-vector refer to the lepton current), the scalar and pseudoscalar operators $\mathcal{O}_S$ and $\mathcal{O}_P$, and the tensor operators $\mathcal{O}_T$ and $\mathcal{O}_{T5}$. Two examples of these global fits are shown as a reference in Fig. 1.6. The plot on the left shows the results of the global fit when two Wilson coefficients, $\mathcal{C}_9^{\mu}$ and $\mathcal{C}_{10}^{\mu}$, are allowed to receive NP contributions. The constraints coming from R_K and R_{K^*} are shown by the solid lines in blue, those coming from $b \to s\mu^+\mu^-$ are shown by the dotted lines in grey, while the combination of both is shown in red, where the different red shades represent the 1σ, 2σ, and 3σ confidence levels. According to the fit, the data are in favour of a reduced $\mathcal{C}_9$ and an unchanged $\mathcal{C}_{10}$ or a reduced $\mathcal{C}_9$ and an enhanced $\mathcal{C}_{10}$. If one allows LFU violation, the Wilson coefficients can assume different values in the electron and muon sectors. The plot on the right shows the results of the global fit when both $\mathcal{C}_9^{e}$ and $\mathcal{C}_9^{\mu}$ are allowed to receive NP contributions. According to LFU, the two couplings should be identical. However, this is not the favoured scenario according to the experimental results, which suggest a suppression in the muon coupling, which is compatible with the suppression observed in the R_K and R_{K^*} measurements. There is instead no clear indication of a NP contribution in the electron sector, given the compatibility of $\mathcal{C}_9^{e}$ with zero.

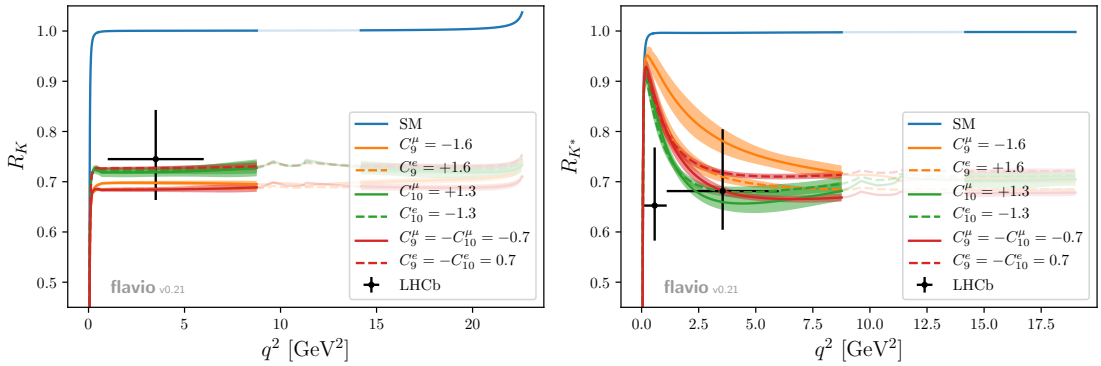


Figure 1.7: Values of R_K (left) and R_{K^*} (right) predicted by the SM (blue line) and by several NP benchmark models (orange, green, and red lines) as a function of q^2, compared to the values measured by LHCb (black points with error bars) [49].

A comparison between the values of R_K and R_{K^*} measured by LHCb and the predictions of the SM and of several NP benchmark models is shown in Fig. 1.7 as a function of q^2. In the SM, R_K and R_{K^*} are to an excellent approximation independent of q^2. For R_K, this is still valid in the presence of NP, although the predicted value of R_K is suppressed with respect to the SM prediction. For R_{K^*}, the q^2 dependence is not trivial and varies with the specific NP model. The drop to zero at the dimuon threshold $q^2 \sim 4m_\mu^2$ is due to phase-space effects.

It is worth noticing that some of the NP models that have been proposed can also explain the deviation observed in the R_{D^*} measurement [50–52]. These models assume the existence of W' and Z' bosons with non-universal couplings to charged leptons. In this models, NP couples predominantly to the third generation of quarks and leptons, with a small non-negligible mixing between second and third generations [53].

1.8 LFU angular asymmetries

The breaking of LFU through NP effects in rare $b \to s\ell^+\ell^-$ transitions can be measured by computing the S_i and P_i' observables for decays to electrons and muons.

The simplest possibility is to compute the angular observables

$$\Delta S_i = S_i^e - S_i^\mu,$$

i.e., the differences between the S_i observables of the same decay to electrons and muons. A more detailed discussion about the expected sensitivity of these observables can be found in 6.1.2.

Another appealing possibility is to compute the angular observables

$$Q_i = P_i'^\mu - P_i'^e,$$

with $i = 4, 5, 6, 8$. These angular observables have a limited sensitivity to hadronic uncertainties and long-distance charm-loop contributions in the SM and can hence be considered a clean probe of LFU violation between the first and the second generation. Any deviation from zero in Q_i would be a direct hint of NP. A pioneering measurement of these angular observables was performed recently

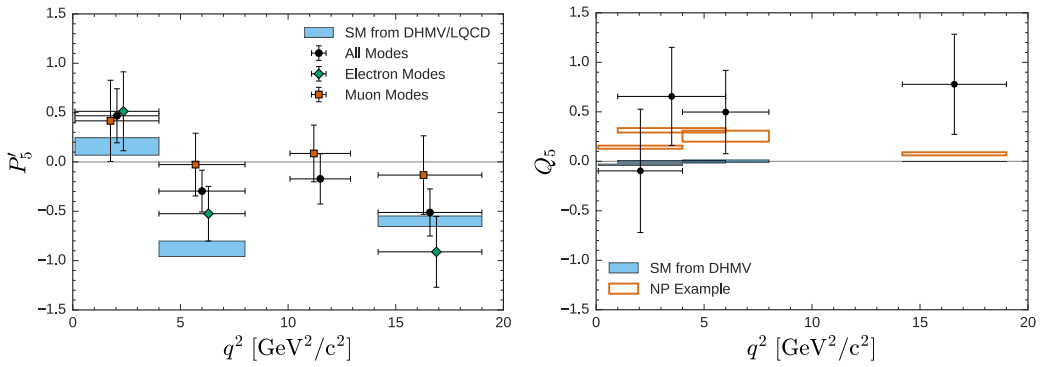


Figure 1.8: Values of P_5' measured by Belle in the electron mode (green), in the muon mode (red), and in the combined mode (black), compared to the SM predictions (blue bands) for different bins of q^2 (left). Values of Q_5 measured by Belle (black points) compared to the SM predictions (blue bands) and to the predictions of a NP benchmark model (orange bands) for the same bins of q^2 (right) [54].

by Belle and is shown in Fig. 1.8. The plot on the left shows the values of P_5' measured by Belle in the electron and muon modes, while the plot on the right shows the corresponding values of Q_5. In order to increase the statistics available, both $B^0 \to K^{*0}\ \ell^+\ell^-$ and $B^+ \to K^{*+}\ \ell^+\ell^-$ decays, with $\ell = e, \mu$, were included in the measurement. The K^{*0} was reconstructed in the $K^+\pi^-$ final state, while the K^{*+} was reconstructed in both $K^+\pi^0$ and $K^0_S\pi^+$ final states. The analysis was performed in four independent bins of q^2.

In order to improve the sensitivity to the difference between the Wilson coefficients $\mathcal{C}_i^e$ and $\mathcal{C}_i^\mu$ for decays to electrons and muons, additional angular observables were proposed recently. One possibility consists in measuring weighted differences (with respect to the lepton flavour of the final state) of angular observables:

$$D_i(q^2) \equiv \frac{d\mathcal{B}^e}{dq^2} S_i^e(q^2) - \frac{d\mathcal{B}^\mu}{dq^2} S_i^\mu(q^2),$$

where $\mathcal{B}^e$ and $\mathcal{B}^\mu$ are the branching ratios of the $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays, respectively [55].

1.9 Summary and future prospects

Rare decays proceeding through a $b \to s\ell^+\ell^-$ transition are a powerful tool to search for NP effects and have shown several hints of tension with respect to the SM predictions in recent years. Tensions have been observed in the LHCb measurements of the LFU ratios R_K and R_{K^*}, the branching ratios of several $b \to s\mu^+\mu^-$ decays, and the P_5' observable of the $B^0 \to K^{*0}\mu^+\mu^-$ decay. The latter has been confirmed by an independent measurement performed by Belle in a very different experimental setting.

This could suggest a coherent pattern of NP effects contributing predominantly to $b \to s\mu^+\mu^-$ and to a lesser extent to $b \to se^+e^-$ transitions. The experimental results might be explained consistently within an effective field theory by assuming NP contributions to one or more Wilson coefficients. Based on current measurements, one of the favoured option would be a NP contribution to $\mathcal{C}_9^\mu$ or to both $\mathcal{C}_9^\mu$ and $\mathcal{C}_{10}^\mu$. This would imply LFU violation between electrons and muons. On the other hand,

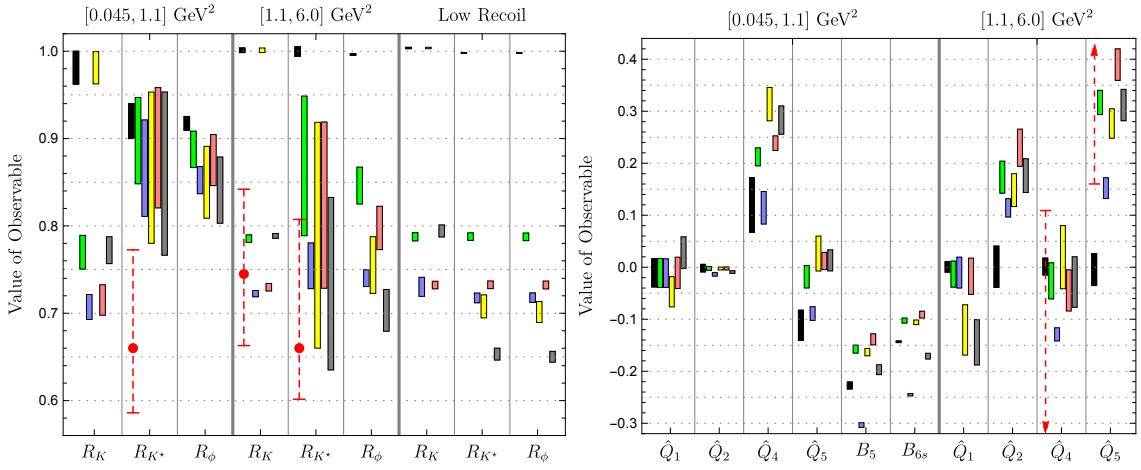


Figure 1.9: Predictions for the ratios R_K, R_{K^*}, and R_ϕ (left) and for the angular observables $\hat{Q}_1$, $\hat{Q}_2$, $\hat{Q}_4$, $\hat{Q}_5$, B_5, and B_{6s} (right) in the low, central, and high q^2 bins for the SM (black) and five NP benchmark models. The latter correspond to $\mathcal{C}_{9\mu}^{NP} = -1.1$ (green), $\mathcal{C}_{9\mu}^{NP} = -\mathcal{C}_{10\mu}^{NP} = -0.61$ (blue), $\mathcal{C}_{9\mu}^{NP} = -\mathcal{C}'_{9\mu} = -1.01$ (yellow), $\mathcal{C}_{9\mu}^{NP} = -3\mathcal{C}_{9e}^{NP} = -1.06$ (orange), and the best fit point in a six-dimensional fit of the Wilson coefficients $\mathcal{C}_7^{NP}$, $\mathcal{C}_{9\mu}^{NP}$, $\mathcal{C}_{10\mu}^{NP}$, $\mathcal{C}'_7$, $\mathcal{C}'_{9\mu}$, and $\mathcal{C}'_{10\mu}$ (grey). The experimental results, if available, are shown by the dashed red bars [56].

some of the observed tensions might be due to the underestimation of the theoretical uncertainties.

Experimentally, most measurements are dominated by statistical uncertainties. The LHCb measurements, in particular, are based on proton-proton collision data corresponding to an integrated luminosity of 3.0 fb^{-1}, recorded at centre-of-mass energies of 7 and 8 TeV. This corresponds to the full statistics of the LHC Run I. However, due to the increase of the centre-of-mass energy during the LHC Run II, nearly twice this integrated luminosity is expected by the end of the current data-taking period. Therefore, it will be a matter of utmost importance to provide updates based on the full LHC Run I and Run II statistics. In addition, new measurements must be performed to further exploit the potential of the rare decays proceeding through a $b \to s\ell^+\ell^-$ transition. Besides R_K and R_{K^*}, additional LFU ratios that can be meaured at LHCb are R_X and R_ϕ, which refer to the decays $B \to X_s\ell^+\ell^-$ and $B_s^0 \to \phi\ \ell^+\ell^-$, respectively. A complementary test of LFU consists in studying the angular distribution of the final-state particles of the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays. This can be achieved by measuring the S_i and P_i' observables of such decays, as well as their differences ΔS_i and Q_i and weighted differences D_i with respect to the flavour of the leptons in the final state.

In order to investigate the potential of these measurements in discovering NP, LFU ratios and angular observables are predicted within several NP benchmark models and compared to the SM predictions and to the measurements (when available). This is shown in Fig. 1.9, where five NP benchmark models are taken into account: in the first, $\mathcal{C}_{9\mu}^{NP} = -1.1$; in the second, $\mathcal{C}_{9\mu}^{NP} = -\mathcal{C}_{10\mu}^{NP} = -0.61$; in the third, $\mathcal{C}_{9\mu}^{NP} = -\mathcal{C}'_{9\mu} = -1.01$; in the fourth, $\mathcal{C}_{9\mu}^{NP} = -3\mathcal{C}_{9e}^{NP} = -1.06$; in the fifth, the Wilson coefficients have the values corresponding to the best fit point in a six-dimensional fit of $\mathcal{C}_7^{NP}$, $\mathcal{C}_{9\mu}^{NP}$, $\mathcal{C}_{10\mu}^{NP}$, $\mathcal{C}'_7$, $\mathcal{C}'_{9\mu}$, and $\mathcal{C}'_{10\mu}$.

At LHCb, the D_i observables can be measured by performing likelihood fits of the mass and angular distributions of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decays or by using the method of moments. However, obtaining an evidence of NP effects at the level of three standard deviations would require more than 1000 $B^0 \to K^{*0}e^+e^-$ decays, which is beyond the reach of LHC Run II. The statistics available

is also the limiting factor of the measurement of the Q_i observables. Due to this limiting factor, a first milestone in the LHCb physics programme is the measurement of the ΔS_i observables. The first measurement of these observables, obtained by a counting method procedure, is described in detail in Chapter 6.

CHAPTER 2

The LHCb experiment at the LHC

This chapter provides some generalities about the Large Hadron Collider and a more detailed description of the LHCb detector.

2.1 The LHC

The Large Hadron Collider (LHC) is a proton-proton (pp) and heavy ion collider located at the European Organization for Nuclear Research (CERN) laboratory, on the French-Swiss border just outside Geneva. The LHC is housed in a 27 km long, nearly circular tunnel about 100 m underground, the same tunnel that previously housed the Large Electron Positron (LEP) collider.

Before circulating into the LHC, protons are extracted from hydrogen gas and accelerated by a succession of accelerators. Each accelerator increases their energy and injects them into the next accelerator in the sequence, as shown in Fig. 2.1. The Linac 2 linearly accelerates protons up to the energy of 50 MeV. Protons are then transferred to the Booster, where they are accelerated up to 1.4 GeV, and subsequently to the Proton Synchrotron (PS), where they reach the energy of 25 GeV. They are then injected into the Super Proton Synchrotron (SPS), and finally transferred into the LHC at an average energy of 450 GeV.

Proton and ion beams are bent around the circumference of the LHC using NbTi superconducting dipole magnets which, maintained at a temperature of 1.9 K by a liquid helium cooling system, produce a field of 8.3 T. Proton beams are not continuous, but spaced in *bunches* of about 10^{11} protons each. The time separation between two adjacent bunches is a multiple of 25 ns, which corresponds to a bunch-crossing rate of 40 MHz. This is referred to as the "nominal" bunch-crossing rate. The LHC was designed to collide protons with a peak instantaneous luminosity of 10^{34} cm^{-2}s^{-1} at a centre-of-mass energy of 14 TeV. However, the energy was limited to 7 TeV during the 2010 and 2011 operations and to 8 TeV during 2012. This first data-taking period, which is referred to as Run I, was followed by an upgrade of the machine and the detectors which took place from 2012 to 2015 (the Long Shutdown I). The changes applied to the accelerator during the Long Shutdown I allowed to further increase the energy to 13 TeV. A second data-taking period, which is referred to as Run

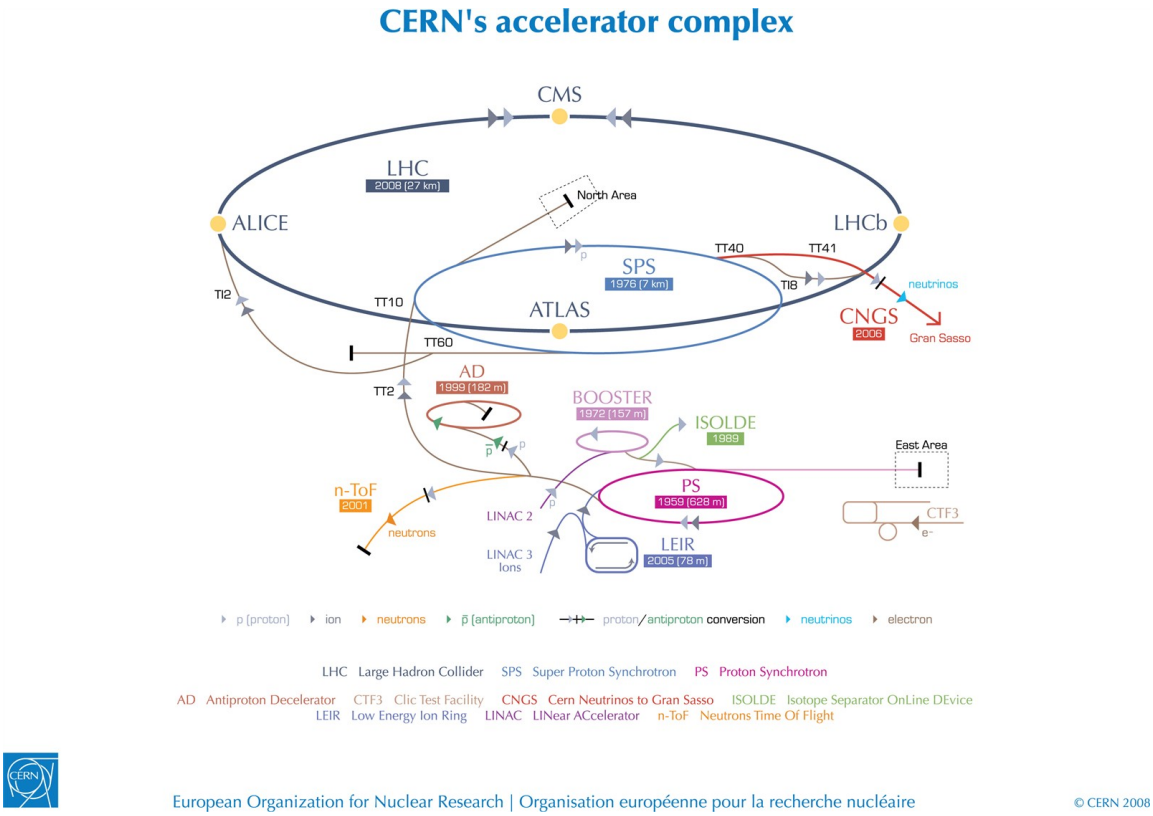


Figure 2.1: Schematic view of the accelerator complex at CERN [57].

II, started in 2015 and is still ongoing. A summary of the main LHC parameters during the 2010, 2011, and 2012 *pp runs* is shown in Table 2.1, while the corresponding summary for the 2015, 2016, and 2017 *pp* runs is shown in Table 2.2.

Collisions between the two beams occur in four distinct points along the ring, where the detectors of the four major LHC experiments (ATLAS, CMS, ALICE, and LHCb) are installed. Among them, ATLAS and CMS are general-purpose experiments, while ALICE and LHCb are specifically designed to study heavy ion and heavy flavour physics, respectively. Two smaller special-purpose experiments, TOTEM and LHCf, complete the picture: the first aims at measuring the total *pp* cross-section, the second aims at studying some aspects of astroparticle physics.

LHC is a copious source of b and c hadrons. At current energies, the total *pp* cross-section is about 100 mb, with the inelastic component amounting to 70 mb [58]. Heavy quark production is dominated by $b\bar{b}$ [59] and $c\bar{c}$ [60] quark-pair production through strong interactions, with cross sections of

$$\sigma(pp \to c\bar{c}X) = (6.10 \pm 0.93)\ \text{mb},$$
$$\sigma(pp \to b\bar{b}X) = (284 \pm 20 \pm 49)\ \mu\text{b},$$

at 7 TeV centre-of-mass energy. An approximately linear increase of the cross sections is expected by raising the energy from 7 to 14 TeV, which will result in nearly 10^5 $b\bar{b}$ and 10^6 $c\bar{c}$ quark-pairs produced every second at 14 TeV centre-of-mass energy. This is consistent with the cross section measurement performed at LHCb at a centre-of-mass energy of 13 TeV [61].

Parameter	Design	2010	2011	2012
Peak luminosity ($\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	1×10^{34}	2.1×10^{32}	3.7×10^{33}	7.7×10^{33}
Protons energy (TeV)	7	3.5	3.5	4
Max bunches per beam	2808	368	1380	1380
Bunch separation (ns)	25	150	50	$25 - 50$

Table 2.1: Main LHC parameters during the 2010, 2011, and 2012 *pp* runs compared with the corresponding design values.

Parameter	Design	2015	2016	2017
Peak luminosity ($\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	1×10^{34}	5.0×10^{33}	13.8×10^{33}	17.4×10^{33}
Protons energy (TeV)	7	6.5	6.5	6.5
Max bunches per beam	2808	2220	2220	2556
Bunch separation (ns)	25	25	25	25

Table 2.2: Main LHC parameters during the 2015, 2016, and 2017 *pp* runs compared with the corresponding design values.

The inclusive $b\bar{b}$ quark-pair production cross-section depends on the rapidities of the b quark and the $\bar{b}$ antiquark and on the transverse mass of the $b\bar{b}$ system [62, 63]. The cross section decreases if either the rapidity difference between the b quark and the $\bar{b}$ antiquark or the transverse mass of the $b\bar{b}$ system increase. This results in $b\bar{b}$ pairs predominantly produced with both partons collinear, as shown in Fig. 2.2. This motivates the design of the LHCb detector, as discussed in the next section.

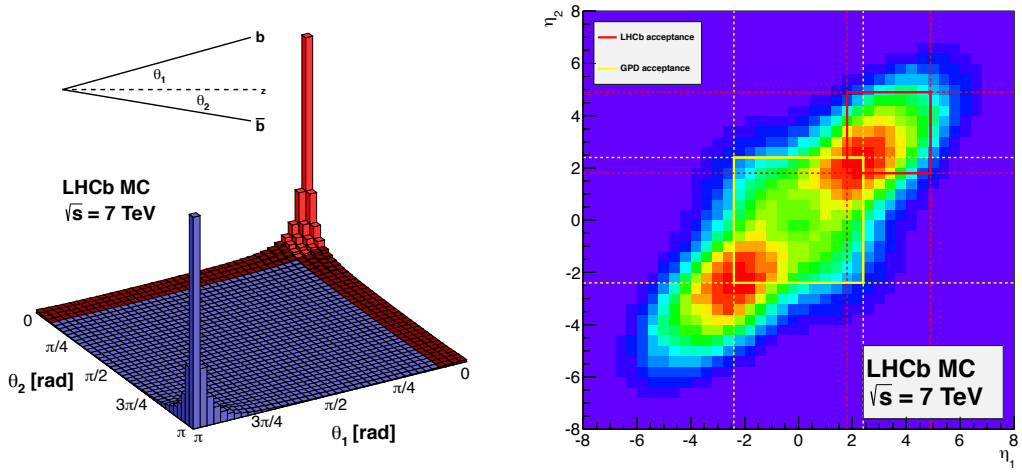


Figure 2.2: Angular correlation between the b quark and the $\bar{b}$ antiquark in $b\bar{b}$ quark-pair production processes [64], as simulated by the PYTHIA8 event generator with the CTEQ6 NLO parton distribution functions.

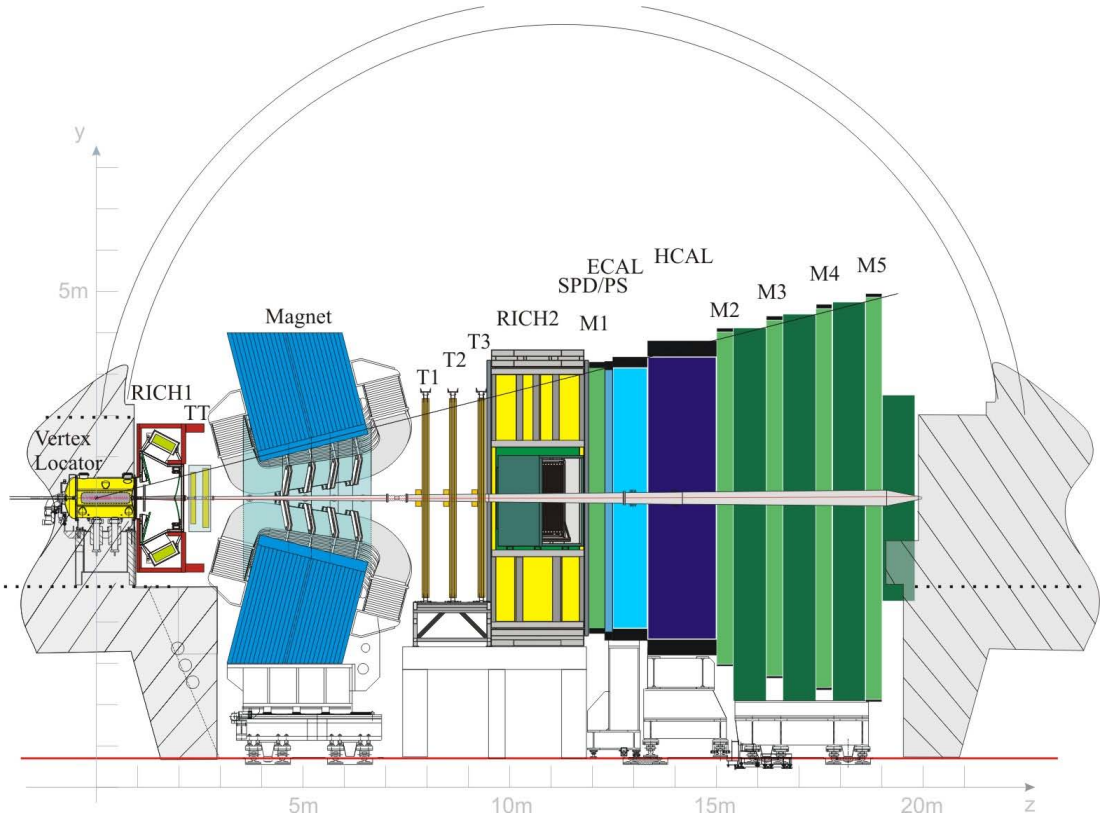


Figure 2.3: Layout of the LHCb detector [65]. The beam is along the z axis.

2.2 The LHCb detector

The primary goal of the Large Hadron Collider beauty (LHCb) experiment is to search for indirect evidences of NP, mainly by studying *CP* violation and rare decays of charm and bottom hadrons.

The LHCb detector [65] is a single-arm spectrometer with a forward angular coverage from 15 mrad to 300 (250) mrad in the bending (non-bending) plane, corresponding to a pseudorapidity interval of $1.8 < \eta < 4.9$, where $\eta = -\ln[\tan(\theta/2)]$, with θ being the polar angle with respect to the beam direction. The forward angular coverage is motivated by the fact that, at the LHC centre-of-mass energies, the b and $\bar{b}$ hadrons that originate from the pp collisions are mostly produced at small polar angles and, once one b quark is in the detector acceptance, probability is high that the corresponding $\bar{b}$ antiquark will be nearby in pseudorapidity. For beam energies of 7 and 8 TeV, the probability that either the b quark or the $\bar{b}$ antiquark are inside the LHCb acceptance is 27%, while the probability for the $b\bar{b}$ pair to be inside the LHCb acceptance is 25%. The corresponding probabilities at 14 TeV are 27% and 24%, respectively [64]. The impact of the acceptance on the $b\bar{b}$ and $c\bar{c}$ quark-pair production cross-sections can be evaluated by comparing the cross-sections quoted in the previous section with the analogous cross-sections for decay products within the LHCb acceptance, which are $\sigma(pp \to c\bar{c}X) = (1419 \pm 134)$ µb and $\sigma(pp \to b\bar{b}X) = (75 \pm 14)$ µb.

The LHCb detector consists of a charged-particle tracking system and a particle-identification system. The tracking system includes a magnet and three different detectors: the VErtex LOcator (VELO) and the Tracker Turicensis (TT), both upstream of the magnet, and three tracking stations (T1-T3), downstream of the magnet. The particle-identification system includes several detectors,

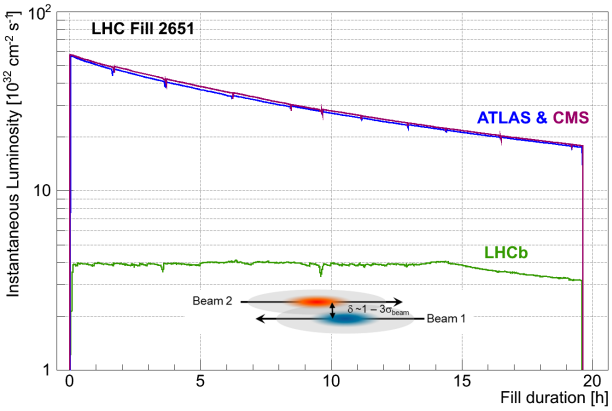


Figure 2.4: Comparison between the instantaneous luminosities at LHCb and at general-purpose detectors for a typical fill duration during the 2010-2012 data taking illustrating the effect of luminosity leveling [67].

each one exploiting a different technology: two Ring Imaging CHerenkov (RICH) detectors, the calorimeters, and the muon detectors. The layout of the LHCb detector is shown in Fig. 2.3. The right-handed coordinate system has the x axis pointing toward the centre of the LHC ring, the y axis pointing upwards, and the z axis pointing along the beam direction.

The LHCb detector guarantees a good coverage of the high pseudorapidity region. However, high particle densities challenge the detectors near the beam axis with high occupancy. This is partially alleviated by using a light-weight beampipe that includes four conical sections that pass through the whole detector. The first three sections, from the interaction point to the calorimeters, are made of beryllium, while the last section, from the calorimeters to the muon detectors, is made of stainless steel.

2.2.1 The collision

The two proton beams, which run parallel to each other along the LHC tunnel, are bent to intersect in the VELO. Primary pp interactions occur in the overlap area, which is referred to as the luminous region. The position and size of the luminous region are inferred from the spatial distribution of the *Primary Vertices* (PV), which are the space-points in which primary pp interactions occur. The size of the luminous region is inferred as the per-run spread of the reconstructed primary vertices x, y, and z coordinates and is approximately 40 µm along x and y and 5 cm along z.

When the beams intersect, multiple primary pp interactions may occur. This challenges the detector with high occupancy and must be limited due to hardware limitations in order to keep the data acquisition efficiency optimal. For this reason, the luminosity at the intersection point was reduced to nearly 3.5×10^{32} cm^{-2}s^{-1} in 2011 and to nearly 4×10^{32} cm^{-2}s^{-1} since 2012. This was achieved by a luminosity leveling technique [66], which allows the LHCb detector to meet occupancy requirements and to operate at the same time as ATLAS and CMS. The transverse overlap of the beams at LHCb is minimum at the beginning of a fill and is then gradually increased to compensate for the beam current degradation, so that the instantaneous luminosity can be kept stable to within nearly 5% during the entire fill. The effect of this procedure is shown in Fig. 2.4, where the instantaneous luminosities at LHCb and at the general-purpose detectors ATLAS and

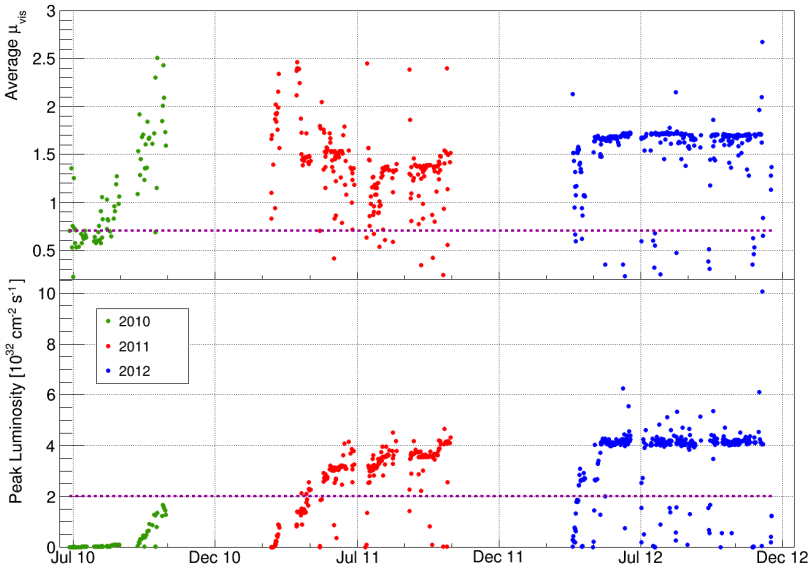


Figure 2.5: Values of μ (top) and peak luminosity (bottom) during Run I, separately for 2010 (green), 2011 (red), and 2012 (blue) [67]. The design values are represented by the dotted violet lines.

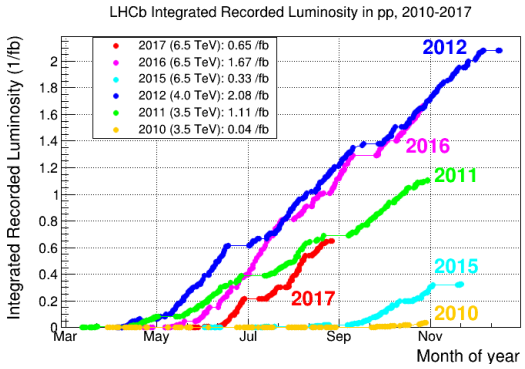


Figure 2.6: LHCb integrated luminosity during *pp* runs as a function of time for 2010 (yellow), 2011 (green), 2012 (blue), 2015 (cyan), 2016 (magenta), and 2017 (red) [69].

CMS are compared for a typical fill duration during the 2010-2012 data taking. The luminosity leveling guarantees similar data-taking conditions during the fill and allows to reduce systematic uncertainties originating from changes in the detector occupancy. The transverse beam size at the intersection point is around 160 µm and the distance between beam centres varies from 100 µm at the beginning of the fill to 40 µm at the end of it. In these conditions, the average number of visible primary *pp* interactions per bunch crossing, called μ, almost reduces to one [68].[1] The values of μ and peak luminosity during Run I are shown in Fig. 2.5. The LHCb integrated luminosity during *pp* runs as a function of time for 2010, 2011, 2012, 2015, 2016, and 2017 is shown in Fig. 2.6 and sums up to 3 fb^{-1} for Run I and 2.6 fb^{-1} for Run II.

Particles that originate from a primary vertex and have significant lifetime and sufficient momentum may travel a measurable distance before decaying. The space-points in which such decays occur

[1] An interaction is "visible" if at least two charged particles with sufficient hits in the tracking system to allow reconstruction are produced.

are referred to as *Secondary Vertices* (SV). Typical c and b hadrons in LHCb have momenta of $\mathcal{O}$ (100) GeV/c and decay lengths of $\mathcal{O}$ (1) cm.

Due to the forward geometry of the LHCb detector, only particles moving in the forward direction may be detected in both tracking and particle-identification detectors. Backward-moving particles may traverse only the most upstream section of the VELO. They are used in the *trigger* to determine the number and the longitudinal positions of the primary vertices.

2.2.2 Tracking

The tracking system provides accurate spatial measurements of charged-particles' trajectories, in order to determine quantities such as charge, momentum, vertex locations, flight distance, and impact parameter (described in more detail below). Moreover, reconstructed tracks are matched to calorimeter clusters, Cherenkov rings, and tracks in the muon detectors (muon stubs), hence providing useful information to discriminate among different species of particles.

The warm dipole magnet allows the measurement of charged-particles' momentum by bending the charged particles in the horizontal plane. The dominant component of the magnetic field is aligned

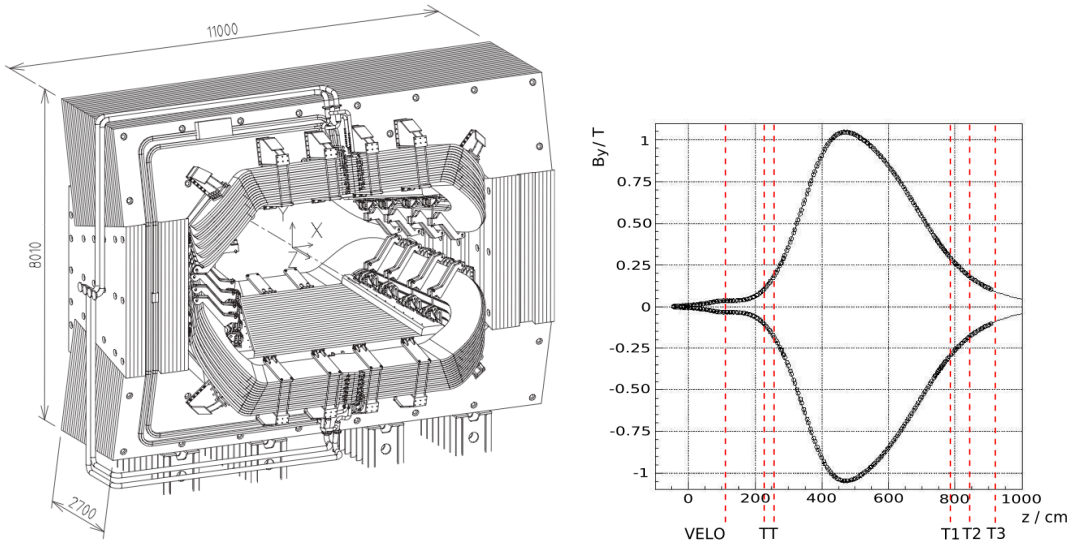


Figure 2.7: Perspective view of the magnet, with dimensions in millimeters and interaction point behind the magnet (left) [65]. Measured on-axis magnetic field, for both "up" and "down" polarities (right) [65].

along the vertical direction and corresponds to 4 T m integrated along 10 m for tracks originating from the interaction point. This guarantees a relative momentum resolution that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. To cover the full acceptance of the LHCb detector, the magnet has an aperture of ±300 (250) mrad in the bending (non-bending) plane. It consists of two identical conical saddle-shape non-superconducting coils, mounted symmetrically inside an iron yoke. Each coil is made of 15 *pancakes*, while the iron yoke is made from 100 mm thick steel sheets. The magnet is 11 m wide, 8 m high, and 5 m thick, and has a total weight of 1600 tons. Its electric power dissipation amounts to 4.2 MW and corresponds to a nominal current of 5.85 kA and a maximum current of 6.6 kA in the conductor. The magnet perspective view is shown in Fig. 2.7. The current in the magnet is periodically inverted and similar amounts of experimental data are collected with each polarity configuration. This reduces the impact of systematic effects in precision measurements

of *CP* asymmetries. The measured on-axis magnetic field is shown in Fig. 2.7, for both "up" and "down" polarities.

The VELO and the TT, upstream of the magnet, are silicon microstrip detectors. The T1-T3 stations, downstream of the magnet, consist of two separated detectors, one closer and one further from the beam axis, that employ a different technology: the Inner Tracker (IT) is a silicon microstrip detector, the Outer Tracker (OT) is a straw tube detector.

VELO

The VELO [70] measures charged-particles' trajectories in the region closest to the interaction point. It is used to reconstruct primary vertices and displaced secondary vertices with a very high spatial resolution. In addition, it allows to precisely measure the Impact Parameter (IP) of charged particles' trajectories. The impact parameter is the distance of closest approach of the particle's trajectory to the primary vertex. Particles originating from the primary *pp* interaction have impact parameter values compatible with zero, while decay products of long-lived heavy particles have higher mean-valued impact parameters. Selections based on impact parameter are extensively used to reduce the contamination from light-quark backgrounds in many measurements at LHCb, so an optimal impact parameter resolution is crucial for the performance of the experiment. The impact parameter of a decay product can be written as

$$\mathrm{IP} = |\vec{x}_{\mathrm{SV}} - \vec{x}_{\mathrm{PV}}| \sin\theta = \frac{|\vec{p}|\tau}{m} \sin\theta,$$

where $|\vec{x}_{\mathrm{PV}} - \vec{x}_{\mathrm{SV}}|$ is the Flight Distance (FD) of the decaying particle, which corresponds to the displacement between the primary and secondary vertices, and θ is the opening angle, in the laboratory frame, of the decay product's trajectory with respect to the decaying particle's trajectory. The opening angle is defined as

$$\cos\theta = \frac{\vec{q}\cdot\vec{p}}{|\vec{q}||\vec{p}|},$$

where $\vec{q}$ and $\vec{p}$ are the decay-product's and decaying particle's momenta, respectively, both measured in the laboratory frame. A schematic view of a typical impact parameter is shown in Fig. 2.8.

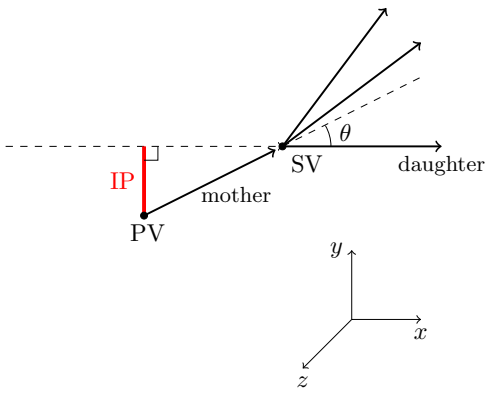


Figure 2.8: Dependence of the impact parameter of a decay product on the flight distance of the decaying particle and the opening angle of the decay product's trajectory with respect to the decaying particle's trajectory.

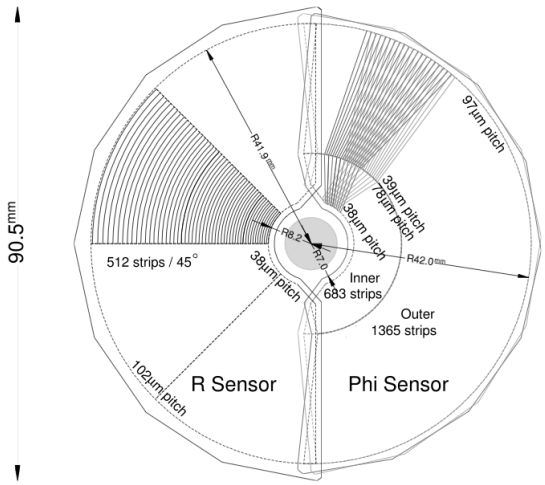


Figure 2.9: Geometry of the VELO r and ϕ sensors [65]. For the ϕ sensor, strips belonging to two adjacent modules are represented in order to show the stereo angle.

The VELO consists of 21 disk-shaped tracking stations positioned along the beam axis, both upstream and downstream of the nominal interaction point. The former aim at identifying primary vertices and providing a precise measurement of their position. Each tracking station is divided in two retractile halves, referred to as *modules*, each consisting of two silicon microstrip sensors, one with radial and one with azimuthal segmentation. Both r and ϕ sensors are centred around the beam position and have a sensitive area covering the region from $r = 8.2$ to $r = 42$ mm. The r sensor consists of concentric semicircular strips, which are subdivided in four 45° sectors each, to reduce occupancy. The pitch increases linearly from 38 µm at the innermost radius to 102 µm at the outermost radius. The ϕ sensor is subdivided in two concentric regions: the inner region at $r = 8.2 - 17.25$ mm, the outer region at $r = 17.25 - 42$ mm. The pitch increases linearly from the centre, with a discontinuity in passing from the inner to the outer region. In order to improve pattern recognition, ϕ sensors are designed with an angular tilt of $+10°$ in the inner region and $-20°$ in the outer region, both with respect to the radial direction. The tilt is reversed for adiacent sensors. The detailed geometry of r and ϕ sensors is shown in Fig. 2.9. The single-hit resolution for tracks of optimal incidence (up to 100 mrad) is measured to be better than 4 µm for a strip pitch of 40 µm. To maximize the impact parameter resolution, r sensors should approach the beam axis as much as possible. Nevertheless, beam dimensions vary considerably between injection and data-taking phases. For detector safety, the two halves are retracted of 29 mm during injection to reach a safe distance from the beam axis. Once stable beam conditions are achieved, the two halves are moved again toward the beam axis until they reach the distance of 5 mm from it, with the sensitive area starting at a distance of 8 mm. Sensors are centred in x and y around the interaction region with an accuracy of 10 µm. The arrangement of the tracking stations is such that a particle originating from the interaction region and emitted in the LHCb acceptance and at a polar angle of less than 390 mrad to the beam axis traverses at least four different layers. Two additional tracking stations, referred to as *veto stations*, are located at the upstream end of the VELO and are assigned to veto pile-up. They are also used to detect interactions between protons and residual gas molecules in the beampipe. The veto stations differ from the other tracking stations because they only use radial

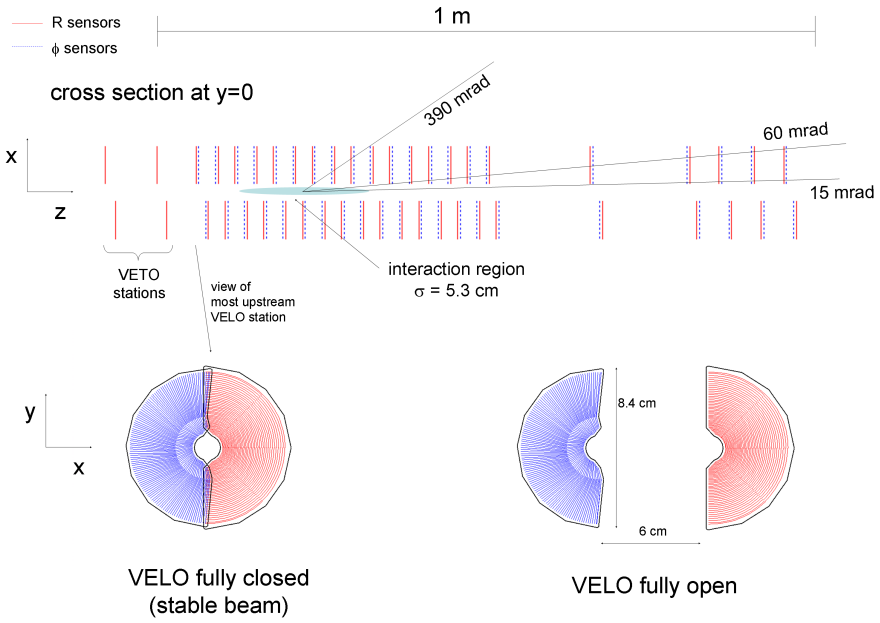


Figure 2.10: Layout of the VELO and sketch of one tracking station in both open and closed configurations [65].

sensors. The layout of the VELO, together with a sketch of one tracking station in both open and closed configurations, is shown in Fig. 2.10.

To minimize the amount of material encountered by particles traversing the detector, all VELO sensors operate in vacuum. This detector vacuum is separated from the beam vacuum through a thin aluminium layer, referred to as the *RF foil*. This layer protects the beam vacuum from outgassing of the modules, provides a shield against Radio Frequency (RF) pickup from the LHC beams, and suppress wake fields that would otherwise affect the LHC beams. The VELO corresponds to 17.5% of a radiation length, the largest contribution to that coming from the RF foil.

The primary-vertex resolution has been measured to be nearly 13 µm in the transverse plane and nearly 71 µm along the z axis for 25 tracks used in the primary-vertex fit [67]. The dependence on the number of tracks, as obtained from 2012 calibration data, is shown in Fig. 2.11. The impact-parameter resolution, which depends on multiple scattering, single-hit resolution, and primary-vertex resolution, can be approximated as

$$\sigma_{\mathrm{IP}} = \sqrt{\sigma^2_{\mathrm{MS}} + \sigma^2_{\mathrm{hit}} + \sigma^2_{\mathrm{PV}}},$$

where the multiple scattering term depends on the inverse of the transverse momentum p_{T} and the single-hit and primary-vertex resolution terms are constants. The typical resolution on impact parameter at LHCb is

$$\sigma_{\mathrm{IP}} = \left(15 + \frac{29}{p_{\mathrm{T}}[\,\mathrm{GeV}/c]}\right)\ \mu\mathrm{m},$$

as obtained from calibration data for events with only one primary vertex [71].

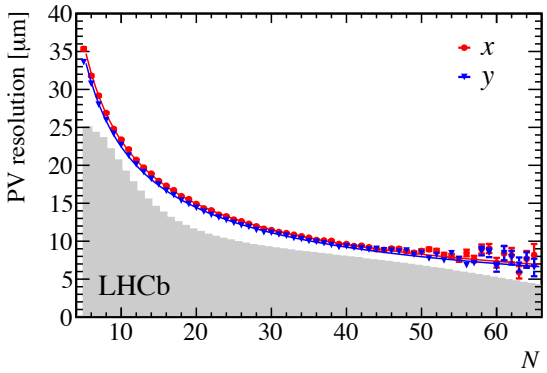


Figure 2.11: Primary-vertex resolution as a function of the track multiplicity, as obtained from 2012 calibration data with only one reconstructed primary vertex in the event [67].

Silicon tracker

The silicon tracker [72] consists of two detectors based on the same technology and sharing the same readout: the TT, upstream of the magnet, and the IT, downstream of it. The TT aims at reconstructing low-momentum tracks that are swept out of the detector acceptance by the magnet and tracks that come from the decay of long-lived particles. The IT reconstructs tracks with momentum larger than 1.5 GeV/c that passed through the magnetic field region and lie near the beam axis. The TT covers the full detector acceptance, the IT covers approximately 2% of the detector acceptance, although it is estimated from simulation that almost 20% of tracks pass through it.

Both detectors are equipped with p^+-on-n silicon microstrip sensors to cope with the high occupancy of the region close to the beam axis. The strip pitch is 183 μm for the TT and 198 μm for the IT, yielding a single-hit resolution of nearly 50 μm. The TT sensors are 9.64 cm wide, 9.44 cm long, and 500 μm thick, while those of the IT are 7.6 cm wide, 11 cm long, and either 320 or 410 μm thick. The TT consists of a single tracking station that has a 150 cm wide and 130 cm high rectangular shape and a nearly 8 m^2 active area, the IT consists of three tracking stations that have a 120 cm wide and 40 cm high cross shape and a nearly 4 m^2 active area, staggered along the beam direction. Each of the four tracking stations is made of four layers, according to a (x,u,v,x) configuration: the outer layers measure the x coordinate with vertical strips, while the second and the third ones have strips rotated $+5°$ (u) and $-5°$ (v) with respect to the vertical direction, to allow a coarse determination of the y coordinate. The distance between two adjacent layers is approximately 4 cm, except for the u and v layers of the TT, which are separated by approximately 27 cm. The layout of both TT and IT layers is shown in Fig. 2.12. Each TT layer consists of half-modules, which have half the TT height and are made of seven silicon sensors each. Depending on the proximity to the beam axis, half-modules are divided in two or three readout sectors, each sector being connected to a readout hybrid. A sector can consists of one, two, three, or four sensors bonded together, which are read out by the same readout hybrid. The IT layer consists of four rectangular units positioned around the beampipe. Each unit is made of seven modules, with each module connected to a readout hybrid: the modules above and below the beampipe contain only one silicon sensor (with a thickness of 320 μm), the modules on the left and on the right contain two silicon sensors (with a thickness of 410 μm) bonded together. In total, the TT and the IT consist of

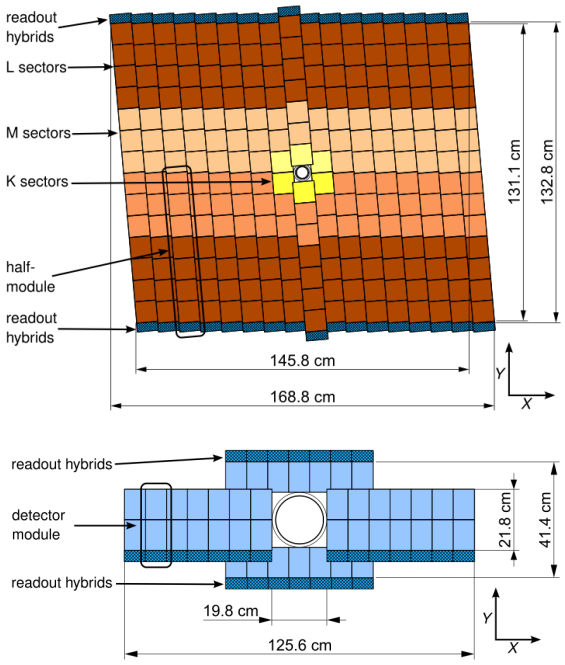


Figure 2.12: Layout of one TT (top) and one IT (bottom) layer [65].

280 readout sectors with 512 strips and 336 readout sectors with 384 strips, respectively.

During Run I, the fraction of working channels was determined to be 99.7% for the TT and 98.6% for the IT. Another key parameter in evaluating the performances of the silicon tracker is the hit efficiency, that is, the ratio between the number of hits found and the number of hits expected in a given sector. During Run I, this was evaluated to be 99.7% for the TT and 99.8% for the IT. The hit resolution is another crucial parameter to take into account. This is determined from the spread of the residuals between the measured hit position and the extrapolated track position. In order to obtain unbiased residuals, the hit is removed from the track before performing the track fit. The hit resolution in 2011 was measured to be 52.6 µm for the TT and 50.3 µm for the IT [67]. For 2012, the hit resolution was measured to be 53.4 µm for the TT and 54.9 µm for the IT [67]. The average occupancy depends on the distance from the beam axis and varies from 1.9% to 0.2% for both TT and IT.

OT

The OT [73] completes the LHCb tracking system, providing a coverage of the detector acceptance up to 300 (250) mrad in the bending (non-bending) plane. The OT uses gas-tight straw tubes to reconstruct tracks with a spatial resolution of 200 µm. Charged particles traversing the detector ionise the gas along their trajectory. By measuring the drift time of the ionisation electrons to the anode wire located at the centre of a straw tube with respect to the bunch crossing signal it is possible to determine the distance between the charged particle's trajectory and the wire.

The OT has the same layout as the IT, as shown in Fig. 2.13. It consists of three tracking stations positioned along the beam axis, each being made of four layers according to the (x,u,v,x) configuration. The total active area of a station is 597 cm $\times$485 cm. A single layer consists of an array of modules, each containing two planes of 64 straw tubes. The straw tubes of one plane are

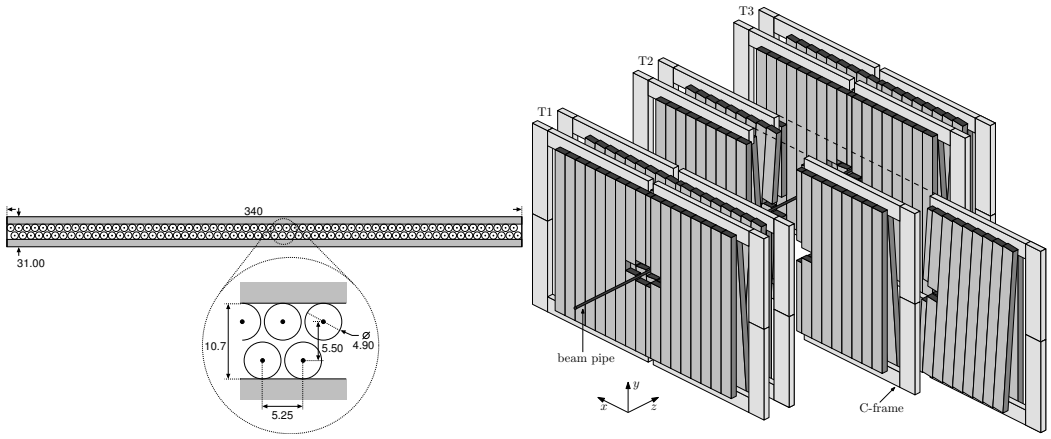


Figure 2.13: Design of an OT module (left) and schematic view of the OT stations along the beam axis (right) [74].

staggered with respect to those of the other plane in order to guarantee overlap between the two adjacent planes. The cathode has a radius of 2.45 mm, the gold-plated tungsten anode wire has a radius of 12.7 µm. Straw tubes are filled with a 70:28.5:1.5 mixture of Ar, CO_2, and O_2 that ensures a drift time across the tube below 50 ns, corresponding to two bunch crossings.

2.2.3 Particle identification

Particle identification plays a crucial role in the analysis of a large fraction of heavy flavour decays studied at LHCb. The variety of charged and neutral particles produced in the collisions, spreading over a wide momentum range, makes an extensive particle-identification system necessary. In particular, efficiency in reducing background often relies on how well the particle-identification system is able to separate kaons from pions, a task that can be accomplished by the RICH detectors. On the other hand, the calorimeters permit identification of electrons, photons, and hadrons, while muons are best identified by the muon detectors. By combining the information of the RICH, calorimeter, and muon detectors, the following particle identification performances are achieved:

- for electrons, 90% identification efficiency for a nearly 5% electron-to-hadron misidentification probability;
- for kaons, 95% identification efficiency for a nearly 5% pion-to-kaon misidentification probability;
- for muons, 97% identification efficiency for a pion-to-muon misidentification probability between 1 and 3%.

More details on the single detectors of the particle-identification system are reported in the following paragraphs.

RICH detectors

Two detectors, RICH1 and RICH2 [75], identify charged particles over a momentum range from 2 to 100 GeV/c. In particular, RICH1 aims to identify low-momentum particles, included those that are swept out of the detector acceptance downstream of the magnet, while RICH2 is optimized to identify high-momentum particles. The low-momentum RICH detector is located upstream of the

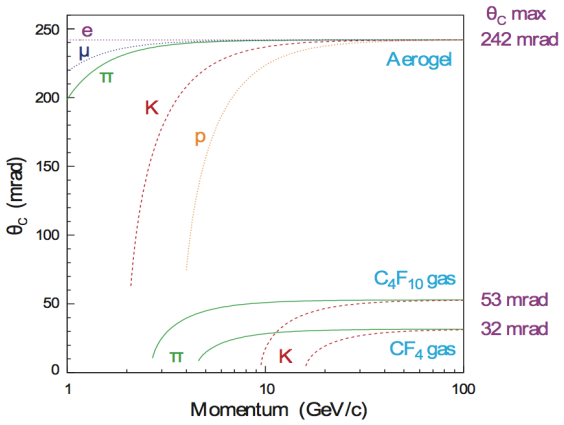


Figure 2.14: Cherenkov angle as a function of momentum for different particles and radiators, as obtained from simulation [65].

magnet, between the VELO and the TT. The high-momentum RICH detector is located downstream of the magnet, just beyond the last tracking station. The relation between Cherenkov angle and momentum vary for each particle and each radiator, as shown in Fig. 2.14. Being designed to cover different momentum ranges, the two detectors differ in the radiators they are filled with. The low-momentum RICH detector covers the range from 2 to 60 GeV/c and uses separate aerogel (in Run I only) and C_4F_{10} radiators. The aerogel has a refractive index of 1.030 at $\lambda = 400$ nm and consists of 5 cm thick tiles placed around the beampipe. The C_4F_{10} has a refractive index of 1.0014 at $\lambda = 400$ nm and fills a 1 m thick area beyond the aerogel. The higher refractive index provides coverage for low-momentum particles. The high-momentum RICH detector covers the range from 15 to 100 GeV/c and uses a CF_4 radiator, which has a refractive index of 1.0005 at $\lambda = 400$ nm. The geometry of both RICH detectors is shown in Fig. 2.15. RICH1 has a wide acceptance, from 25 mrad to 300 (250) mrad in the bending (non-bending) plane, while RICH2 only covers the acceptance from 15 mrad to 120 (100) mrad in the bending (non-bending) plane. Each detector has two kinds of mirrors: a spherical mirror needed for ring-imaging and a set of flat mirrors needed to guide photons onto the Hybrid Photon Detectors (HPDs) located outside the detector acceptance. This minimizes the amount of material traversed by the particles and accommodates the magnetic shield necessary for a proper operation of the hybrid photon detectors. These are used by both RICH detectors and detect Cherenkov photons with wavelengths between 200 and 600 nm.

Calorimeters

The calorimeters provide fast information for the low-level trigger and offer identification of electrons, photons, and hadrons, together with a measurement of their energies and positions.

The calorimeters consist of an Electromagnetic CALorimeter (ECAL) [76], followed by a Hadron CALorimeter (HCAL) [77]. Both are placed between the first and the second muon station and cover the angular acceptance from 25 to 300 (250) mrad in the bending (non-bending) plane. The inner angular acceptance is limited by the high particle density of the region close to the beam axis. The ECAL is equipped with two additional detectors, placed in front of it and separated by a thin lead converter: a PreShower detector (PS) and a Scintillator Pad Detector (SPD). The PS and the SPD are used by the low-level electron trigger to distinguish electrons from photons and pions [78].

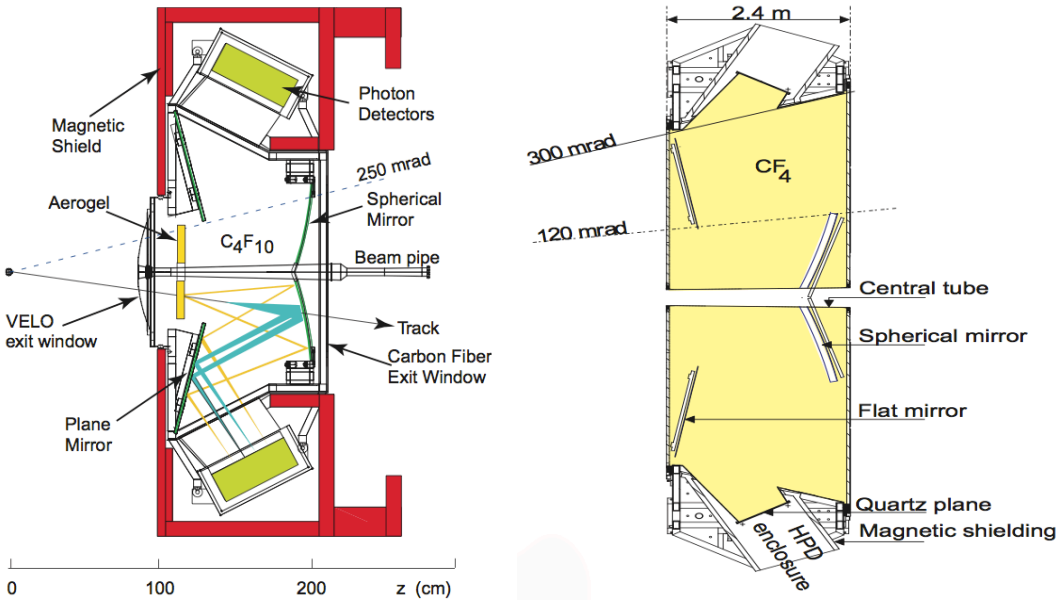


Figure 2.15: Geometry of low-momentum (left) and high-momentum (right) RICH detectors [65]. The optical layout is vertical for the former and horizontal for the latter.

The SPD is also used to measure the number of tracks per event, which is used to veto online too crowded events. The calorimeters have a tower-projective geometry with respect to the nominal interaction point and are subdivided in four quadrants that surround the beampipe. Each quadrant has a lateral segmentation in cells of different sizes, depending on the distance from the beam axis. The lateral segmentation takes into account the variation in hit density of nearly two orders of magnitude over the surface of the calorimeters and is finer in the PS, SPD, and ECAL and coarser in the HCAL, as shown in Fig. 2.16. The PS, SPD, and ECAL consist of three sections, referred to as inner, middle, and outer, while the HCAL consists of two sections, referred to as inner and outer. The thickness of the ECAL corresponds to 25 radiation lengths, to guarantee a nearly complete electromagnetic shower containment and a good energy resolution. The thickness of the HCAL is limited to 5.6 interaction lengths due to space constraints. The readout is common to all detectors: scintillation light is transmitted to photo-multipliers using wavelength-shifting fibers. The front-end electronics, and the PS and SPD photo-multipliers, are located outside the detector acceptance, while the ECAL and HCAL photo-multipliers are placed directly on the detector modules. Both ECAL and HCAL are sampling calorimeters. The ECAL consists of alternate 4 mm thick scintillator tiles and 2 mm thick lead plates. This structure offers a fast response for the low-level trigger and an energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, where the energy is expressed in GeV [65]. The HCAL is structured in 4 mm thick scintillator tiles sandwiched between 16 mm iron sheets. Scintillator tiles run parallel to the beam axis and have a depth that corresponds to one interaction length in iron. Adjacent layers consists of iron sheets alone or paired with scintillator tiles, leading to a ratio of active to passive material of 0.18. This design allows the HCAL to provide a fast response for the low-level trigger, although with a limited energy resolution of $\sigma_E/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$ (E expressed in GeV), as determined from test-beam data [65].

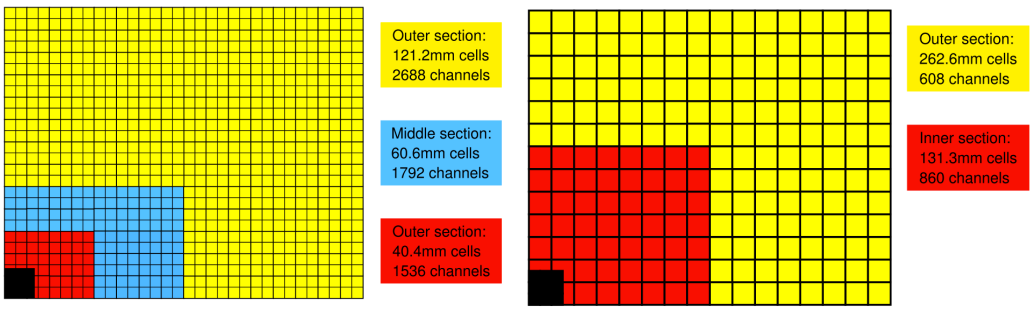


Figure 2.16: Lateral segmentation of PS, SPD, and ECAL (left) and of HCAL (right) [65]. Only one quadrant of the detector is shown. The black area represents the gap for the beampipe.

Muon detectors

Muons in the final state are a typical signature for some of the most interesting heavy flavour decays such as $B^0 \to K^{*0}\mu^+\mu^-$, $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$, $B_s^0 \to \mu^+\mu^-$, and $B^0 \to \mu^+\mu^-$. Moreover, muons play an important role in measurements of B^0 and B_s^0 meson oscillations, where they are used as an efficient tag of the initial flavour of the neutral meson.

The muon detectors [79] provide identification and transverse momentum measurement of penetrating muons for both low-level and high-level triggers, as well as for offline reconstruction. They consist of five stations of rectangular shape, referred to as M1-M5, placed along the beam axis and covering the angular acceptance from 20 (16) to 306 (258) mrad in the bending (non-bending) plane (Fig. 2.17). The station M1, installed between the high-momentum RICH detector and the calorimeters, improves transverse momentum measurements for muons that are also detected in the M2-M5 stations. These stations are placed downstream of the calorimeters. They are interleaved with 80 cm thick iron absorbers that select penetrating muons and result in a total thickness of about 20 interaction lengths. In order to traverse the whole detector, a muon is typically required to have at minimum momentum of 6 GeV/c. The geometrical arrangement of the stations is projective and pointing to the nominal interaction point. The stations are subdivided in four quadrants, arranged around the beampipe. As shown in Fig. 2.17, each quadrant comprises four regions, labelled with R1-R4 and installed at increasing radii from the beampipe. Each region is segmented in pads, whose size depends on the station and on their location within the station. The linear dimensions of the R1, R2, R3, and R4 regions scale according to the ratio 1:2:4:8, so that each region has approximately the same occupancy as the others. The muon detectors rely on two technologies to detect muons: triple gas electron multiplier and multi-wire proportional chamber detectors. The former are used in the innermost region (R1) of the first station (M1), where the high particle density requires a radiation tolerant detector; the latter are used in the rest of detectors. The gas mixture consists of Ar, CO_2, and CF_4 for both detectors, although in different proportions. The first three stations (M1-M3) contribute to transverse momentum measurements, while the last two stations (M4-M5) achieve the simpler task of detecting particles that pass through the absorber material. An average transverse momentum resolution of 20% is achieved in stand-alone muon reconstruction, which is used in the trigger.

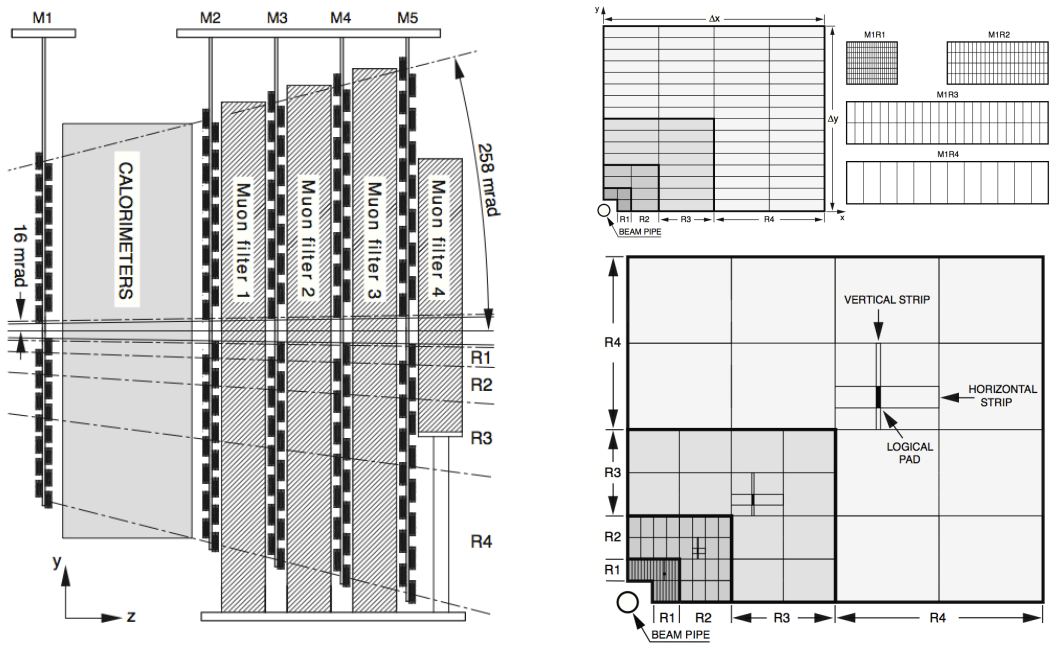


Figure 2.17: Side view of the muon detectors (left) and geometry of a quadrant, with each rectangle representing one chamber (right) [65]. Linear dimensions and segmentations of R1, R2, R3, and R4 scale in the ratio 1:2:4:8.

2.2.4 Trigger

The LHCb trigger [80] is designed to efficiently select heavy flavour decays from the copious light-quark background. The trigger must sustain the 40 MHz bunch-crossing rate, of which nearly 11 MHz consist of crossings where at least one visible primary pp interaction occurs. Only a small fraction of approximately 15 kHz contains a b hadron decay for which all the final-state particles are emitted in the detector acceptance. The subset of interesting b hadron decays is even smaller, corresponding to only few Hertz [65]. The maximum rate at which events can be stored at LHCb varies from 3.5 to 12.5 kHz. It is then crucial for the trigger to reject background as early as possible in the data flow. This is achieved by using fast, coarse measurements of transverse momentum, transverse energy, track displacement, muon identification, and topological properties which are distinctive of some specific decays.

The trigger is organized in two levels: the Level-0 trigger (L0) and the High-Level Trigger (HLT). This two-level structure allows coping with timing and selection requirements, with a fast and partial reconstruction at low level, followed by a more accurate and complex reconstruction at high level. The hardware-based L0 trigger operates synchronously with the bunch crossing. It uses information from calorimeters and muon detectors to reduce the 40 MHz bunch-crossing rate to below 1.1 MHz, which is the maximum value at which the detector can be read out by design. After that, the asynchronous software-based HLT trigger performs a finer selection based on information from all detectors and reduces the rate to an output of up to 12.5 kHz.

The L0 trigger

The L0 trigger consists of three independent trigger decisions: the L0 pileup, the L0 muon, and the L0 calorimeter. Each decision is combined with the others through a logic "or" in the L0 decision unit, reducing the 40 MHz bunch-crossing rate to below 1.1 MHz. The L0 decision unit provides the global L0 trigger decision, which is transferred to the readout supervisor board and, subsequently, to the front-end boards. This is necessary since the full detector information for a given bunch crossing is not read out from the front-end boards until the L0 decision unit has accepted it. Data from all detectors are stored in memory buffers consisting in an analogue pipeline that is read out with a fixed latency of 4 μs, within which time a trigger decision must be available. To accomplish this task, the L0 trigger is entirely based on custom-built electronic boards. At this stage, trigger requests can only involve simple and immediately available quantities, like those provided by calorimeters and muon detectors.

The L0 pileup trigger contributes to luminosity measurements [81].

The L0 muon trigger [82] uses the information from the five muon stations to identify the most energetic muons. It employs four independent processors, each of them associated to a quadrant of the muon detectors. Muons traversing more than one quadrant are not used for triggering. Each processor looks for hits defining a straight line projecting from the nominal centre of the interaction region and traversing all the stations in the y projection. The search is limited to muons with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}/c$ in the x projection. The first and second stations are then used to coarsely estimate the transverse momentum of the muon candidate. Once the two highest-transverse-momentum muon candidates per quadrant are identified, the trigger decision is set depending on two thresholds: one on the highest transverse momentum and one on the product of the two highest transverse momenta (L0 muon and L0 dimuon, respectively). Typical thresholds in 2011 (2012) were $1.48\,\mathrm{GeV}/c$ ($1.76\,\mathrm{GeV}/c$) on the highest transverse momentum and $(1.30\,\mathrm{GeV}/c)^2$ ($(1.60\,\mathrm{GeV}/c)^2$) on the product of the two highest transverse momenta. The number of hits in the SPD is used to veto events with high charged-particle multiplicity. These events are discarded since they would occupy a disproportionate fraction of the HLT processing time. The threshold on the maximum number of hits in the SPD is fixed to 600 (900) for the L0 muon (L0 dimuon).

The L0 calorimeter trigger [83] uses the information from the ECAL, the HCAL, the PS, and the SPD. It calculates the transverse energy deposited in a cluster of 2×2 cells of the same size, for both the ECAL and the HCAL. The transverse energy is defined as

$$E_{\mathrm{T}} = \sum_{i=1}^{4} \mathrm{E}_i \sin\theta_i,$$

where E_i is the energy deposited in cell i and θ_i is the angle between the beam axis and the direction of the particle, assumed to originate from the centre of the interaction region and to hit the centre of cell i. This quantity is combined with information on the number of hits in the PS and SPD to define three types of trigger candidates, photon, electron, and hadron. The photon candidate (L0 photon) is associated with the highest transverse-energy cluster in the ECAL, provided that hits are present in the PS and absent in the SPD, as expected for neutral particles. The definition of a photon candidate differs if the cluster is identified in the inner region of the ECAL. In this case, a cluster accompanied by a suitable number of hits in the PS is sufficient. Electron candidates (L0 electron) are defined similarly to photon candidates, the only difference being the additional requirement of

hits in the SPD. The hadron candidate (L0 hadron) is associated with the highest transverse-energy cluster in the HCAL. If the highest transverse-energy cluster in the ECAL is geometrically matched to it, the two transverse energies are summed to get the total transverse energy of the hadron candidate. The transverse energy of each candidate is compared to a predefined threshold and a positive trigger decision is set for events containing at least one transverse-energy deposit exceeding the threshold. Typical thresholds in 2011 (2012) were 3.5 GeV (3.7 GeV) for hadron candidates and 2.5 GeV (3.0 GeV) for photon and electron candidates. The threshold on the number of hits in the SPD is fixed to 600.

The L0 electron and hadron triggers have a lower efficiency with respect to the L0 muon trigger. This is due to the fact that they both rely on the measurement of transverse energy in the calorimeters, while the L0 muon trigger uses the transverse momentum measured by the muon detectors. The correlation between the measured transverse energy and the transverse momentum of the particle triggering the event is limited in several ways. First of all, the magnetic field kick changes the transverse momentum of charged particles prior to their detection in the calorimeter in a way that depends on their trajectory and charge. Another factor that spoils the $E_{\mathrm{T}} - p_{\mathrm{T}}$ correlation is the calorimeter granularity. The transverse energy measured in a given calorimeter cell is not necessarily deposited by a single charged particle. Cell acceptance is such that transverse energy observed in a cell is the sum of transverse energies released in that cell by all particles traversing it. Especially for central cells, where the particle density is high and low-momentum particles from the beam halo may interfere, several particles contribute to the transverse energy measured in a cell, mimicking the passage of a single energetic particle. The current performance of the hadron trigger is driven by the need to lower the rate of L0 triggered events to 1.1 MHz, which is the maximum rate at which the detector can be read out. This is achieved by requiring a high transverse energy threshold. Such a restrictive requirement is successful in reducing the event-accept rate but causes a loss of a substantial fraction of signal events too. For electrons, an additional complication arises from the emission of *bremsstrahlung* photons, which may or may not hit the same calorimeter cell of the emitting particle. This is discussed in more detail in Chapter 6.

The HLT

Events accepted at L0 are transferred to the Event Filter Farm (EFF), an array of computers consisting in 29000 cores during Run I and more than 50000 cores during Run II.

The HLT consists of several trigger selections designed to collect specific events (*e.g.*, c or b hadron decays). Every trigger selection is specified by reconstruction algorithms and selection criteria that exploit the kinematic features of charged and neutral particles, the decay topology, and the particle identities. Since trigger selections may vary from run to run, a hexadecimal trigger configuration key, which is a list of active trigger selections and associated requirements, is associated to every event, to keep track of trigger biases in subsequent analysis.

The HLT consists of two different levels: the first stage (HLT1) and the second stage (HLT2). The main differences are the complexity of the information these stages are able to process and the available time they have to do so. A partial event reconstruction is done in the first stage in order to significantly reduce the event-accept rate and a more complete event reconstruction follows in the second stage.

At the first level, tracks are reconstructed in the VELO and selected based on their probability

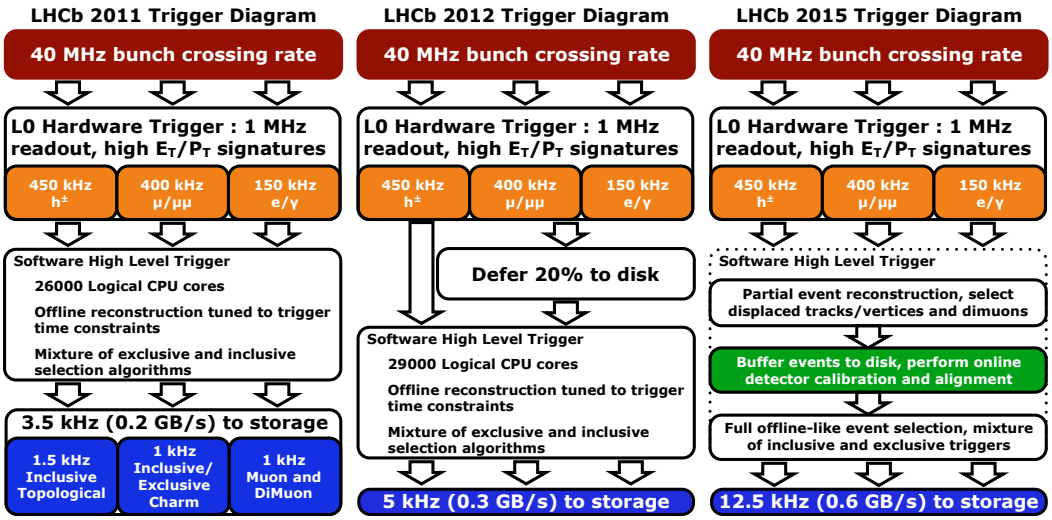


Figure 2.18: Trigger flow of the LHCb experiment during 2011 (top left), 2012 (top right), and 2015 (bottom), illustrating trigger selections and typical event-accept rates after each stage [84].

to originate from heavy flavour decays. This is achieved by determining their impact parameter with respect to the closest primary vertex and their quality, or by identifying the subset of tracks that geometrically match to muon candidates for events triggered by L0 muon or L0 dimuon. Selected tracks are then associated with track segments in the tracking stations (forward tracking), thus allowing measurement of the transverse momentum of the corresponding charged particles. Remaining tracks are ignored by the online pattern recognition algorithm in order to speed up the decision process. This differs from what happens offline, where forward tracking is applied to all tracks in the event reconstructed in the VELO. Once forward tracking has been performed, events are selected by looking at tracks with high momentum and transverse momentum and by selecting on track χ^2 and impact parameter χ^2 [85]. The latter, referred to as χ^2_{IP}, is similar to the impact parameter, but takes the uncertainty on the charged-particle's trajectory into account.

At the second level, a complete forward tracking of all tracks reconstructed in the VELO is performed. While two tracking algorithms are available offline, only one of them (which is based on seeding the search with VELO tracks) is implemented online. Then, tracks are associated with muon and electron candidates, using the offline muon identification algorithm and matching tracks with clusters in the ECAL, respectively. Secondary vertex reconstruction is performed and requirements on decay length and mass are applied. Several trigger selections, either inclusive or exclusive, are available at this stage. Inclusive trigger selections, usually referred to as *topological lines*, trigger on partially-reconstructed b hadron decays. They aim at collecting all b hadron decays with a displaced secondary vertex and at least two charged particles in the final state, thus allowing the selection of events in which only a subset of the decay products is reconstructed. This leads to high signal efficiencies but also to high event-accept rates, which are reduced by imposing additional requirements. To reduce the background due to false tracks that originate from the combination of random hits, all tracks are required to have a track χ^2 per degree of freedom smaller than five, where the number of degrees of freedom depends on the track and corresponds to the difference between the number of track measurements contributing to the track fit and the number of parameters necessary to parametrize the track (four, for a straight-line track, or five, if the curvature due to the

magnetic field is taken into account). To reduce the background due to prompt particles, all tracks are required to have an impact parameter χ^2 greater than 16. All particles are assigned the kaon mass and used to form multibody candidates: two input particles are combined to form a two-body object; another input particle is added to the two-body object to form a three-body object, and so on. At each step, the distance of closest approach between the $(n-1)$-body object and the input particle is required not to exceed 0.15 mm. The multibody candidates are then required to have

$$4 < m_{\rm corr} = \sqrt{m^2 + |p_{\rm Tmiss}|^2} + |p_{\rm Tmiss}| < 7\,{\rm GeV}/c^2,$$

where $m_{\rm corr}$ is a "corrected" mass that includes the contribution of potentially missing decay products. This is achieved by adding the missing momentum transverse to the direction of flight $p_{\rm Tmiss}$ (as determined by primary and secondary vertex positions) to the multibody candidate mass m. Further requirements involve the transverse momentum and the impact parameter χ^2 of decay products and the quality of their tracks. Other trigger selections exploit tracks that are identified as muons. They require single muons to have high transverse momentum (and eventually large impact parameter χ^2). They also use dimuon candidates and require them to have large flight distance and high transverse momentum. Exclusive trigger selections aim at specific b hadron decays. They require all decay products to be reconstructed and impose stringent requirements on quantities such as the candidate mass. Their event-accept rates are usually modest but the associated event samples can hardly be used for channels different from those they target.

Accepted events at the first and second level are finally written to storage. All stored events are then processed offline, using more sophisticated algorithms and adding the latest alignment and calibration information. Offline reconstructed events are divided into separate streams targeted at studying subsets of interesting physics channels, such as the b or c hadron streams, which are the starting point of the physics analysis.

To make a more efficient use of the computing power of the event filter farm, a deferral procedure has been adopted starting from 2012. This consists in deferring a fraction of the HLT processing to the time between two consecutive fills, which is typically several hours. With this procedure, nearly 20% of the events accepted by the L0 trigger are temporarily stored on the nodes of the event filter farm and are further processed by the HLT at a later stage, thus allowing to increase the statistics of the samples used for physics measurements. In addition to the deferred trigger, the HLT workflow has undergone major changes during the Long Shutdown I, due to several improvements in the computing infrastructure. From 2015, the two levels of the HLT have become two independent asynchronous processes running on the same node of the event filter farm. All events passing the first-level trigger are now buffered on the local disk and then processed by the second-level trigger according to a priority system that allows to make a better use of the available resources. During the data taking, the first-level trigger has a high priority, so that more events can pass its selection and be ready to be processed by the second-level trigger, whose priority is increased only when no pp collisions are taking place. The separation between the two levels allows to add the alignment and calibration information before the events are selected by the second-level trigger, thus providing a way to select events with an offline-quality reconstruction already in the trigger. These changes allowed to obtain signal candidates for physics measurements only few hours after the data taking. Additional improvements include lowering the minimum momentum for which a particle can be reconstructed. Typical HLT event-accept rates were 3.5 kHz in 2011 and 5 kHz in 2012, after the

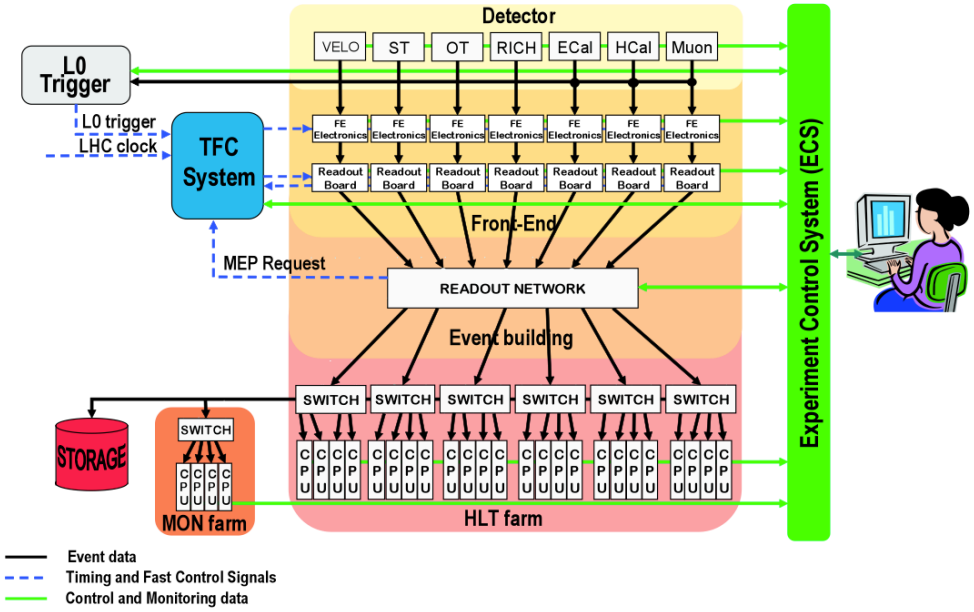


Figure 2.19: Architecture of the online system [65].

deferred trigger had been introduced [80]. In 2015, after the major changes implemented during the Long Shutdown I, rates of up to 12.5 kHz have been achieved. An overview of the trigger flow during 2011, 2012, and 2015 is shown in Fig. 2.18.

2.2.5 Online system

The online system [86] includes all the processes necessary to transfer data from the front-end boards to permanent storage. It consists of three main components: the Data AcQuisition system (DAQ), the Timing and Fast Control system (TFC), and the Experiment Control System (ECS), which are schematically shown in Fig. 2.19. While some information from calorimeter and muon detectors is processed by the L0 trigger, the full detector information is stored in pipelines of the front-end boards, which are detector-specific and custom-made electronic boards. If a positive decision is returned by the L0 trigger, data are moved from the front-end boards to the readout boards, where zero-suppression is performed. The transfer is triggered by the readout supervisor board inside the timing and fast control system. The timing and fast control system, which is synchronous with the LHC clock and the L0 trigger, specifies a destination address for the data produced by every readout board. The destination address identifies one of the computers of the farm, which receives and processes the data from that particular readout board. All fragments of the same event are sent to the same node in the farm for event building and subsequent processing. To reduce readout network overhead, fragments of several consecutive events are packed together in multi-event packets prior to transmission to the readout network. Once the processing is completed, data are sent to storage and an event request (that specifies that the farm is ready to process a new packet of events) is sent to the timing and fast control system. Every step is controlled and supervised by the experiment control system, that monitors the whole online system, the trigger, as well as operation parameters like temperature, pressure, high and low voltages. More details about the readout architecture can be found in Ref. [87].

CHAPTER 3

The LHCb upgrade

This chapter provides an overview of the LHCb upgrade and a more detailed description of the Upstream Tracker, which will substitute the Tracker Turicensis after the Long Shutdown II.

3.1 Motivation

The data collected at LHCb during Run I and those that are being collected during Run II allow to perform precision measurements of rare decays of b and c hadrons in order to test the SM and set new limits on models involving NP contributions. Despite the great success of LHCb, however, several precision measurements will still be limited by statistical uncertainties even at the end of Run II. With the current configuration, several years of data taking would be needed to significantly improve the precision of some key measurements and to approach the uncertainties of the corresponding theoretical predictions. In order to increase the annual event yields, the LHCb detector will undergo a major upgrade during the Long Shutdown II [88,89]. The changes will allow the detector to take data at an increased instantaneous luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (that is, five times higher than in the current configuration) and to read out the full detector at the bunch-crossing rate of 40 MHz. At the instantaneous luminosities foreseen after the Long Shutdown II, the average number of primary pp interactions per bunch crossing will be between 3.8 and 7.6 compared to nearly one in Run II. The integrated luminosity collected by LHCb during Run I and Run II will be nearly $8\,\mathrm{fb}^{-1}$, while the changes planned for the Long Shutdown II will allow LHCb to record $5\,\mathrm{fb}^{-1}$ per year.

3.2 Overview of the LHCb upgrade

The layout of the upgraded LHCb detector is shown in Fig. 3.1.

As discussed in the previous chapter, one of the limitations of the current detector is the low efficiency of the L0 hadron trigger. This is due to the fact that a high E_{T} threshold is needed in order to limit the event-accept rate of the L0 trigger to below 1.1 MHz, which is the maximum rate

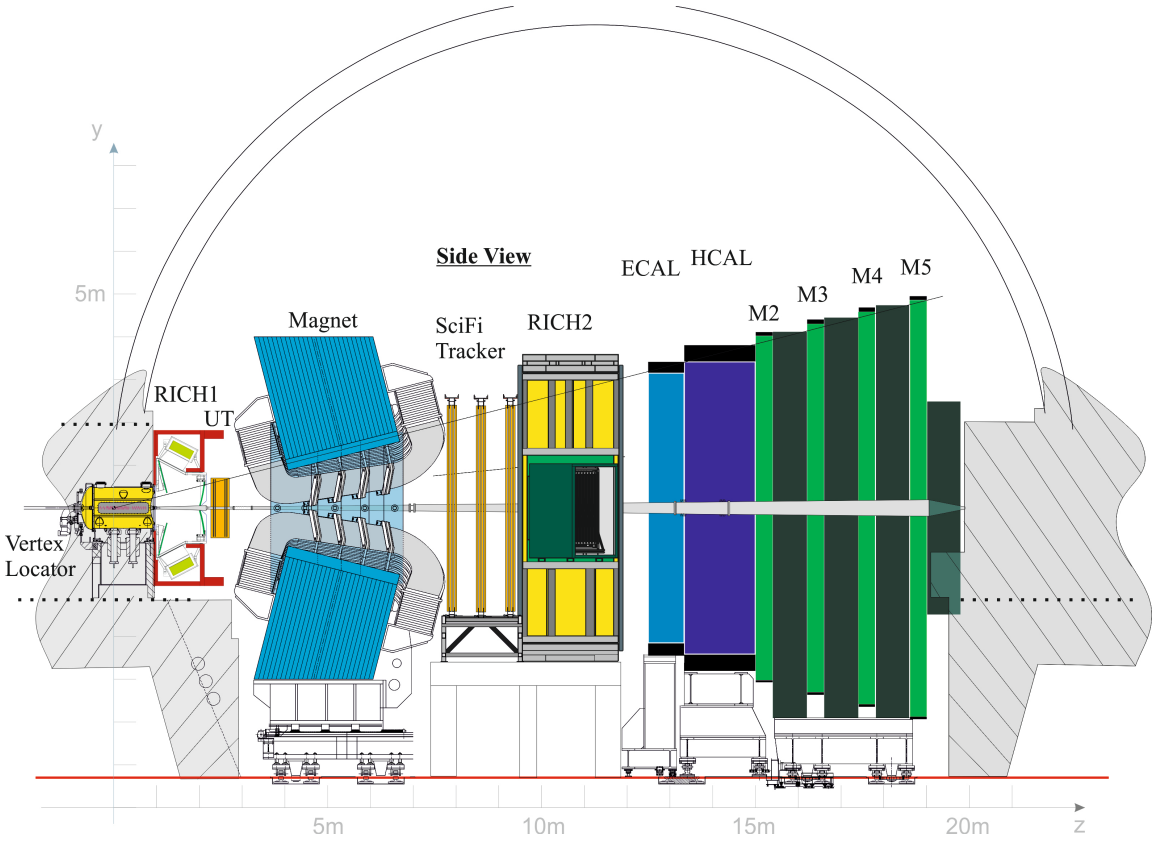


Figure 3.1: Layout of the upgraded LHCb detector [90]. The beam is along the z axis.

at which most subdetectors can be read out. These subdetectors will be redesigned so that the full detector can be read out at the bunch-crossing rate of 40 MHz.

The upgraded VELO detector [91] will adopt pixels instead of microstrips, will have a finer granularity, and will be positioned closer to the beam, with its sensitive area starting at a distance of 5.1 mm from the beam axis. This will allow to improve the reconstruction of vertices and impact parameters and cope with the high occupancy expected in the region closest to the beam. The expected improvement in impact parameter resolution and efficiency with respect to the current VELO detector is shown in Fig. 3.2. The upgraded VELO detector will consist of 41 million 55 µm wide and 55 µm high pixels, which will be read out at the 40 MHz bunch-crossing rate by the custom-built VeloPix front end ASIC [70]. In order to cool the sensors and decrease radiation damage, cope with the power dissipation of the ASIC, and respect the material budget constraints of the detector, the cooling system will be embedded in the modules, by means of evaporative CO_2 circulating through miniature channels into silicon substrates. The layout of a module of the upgraded VELO detector is shown in Fig. 3.3.

The IT and OT will be replaced by the Scintillating Fibre Tracker (SciFi) [90], which will cover the full acceptance downstream of the magnet and guarantee a spatial resolution of nearly 80 µm. The SciFi will consist of three stations of four detection planes according to the (x,u,v,x) configuration. Each detection plane will have five staggered layers of scintillating fibres to ensure a hit efficiency above 99%. The modules will consist of scintillating fibres with a radius of 125 µm and a length of 2.5 m, which will be read out by Silicon PhotoMultipliers (SiPMs) located at the top

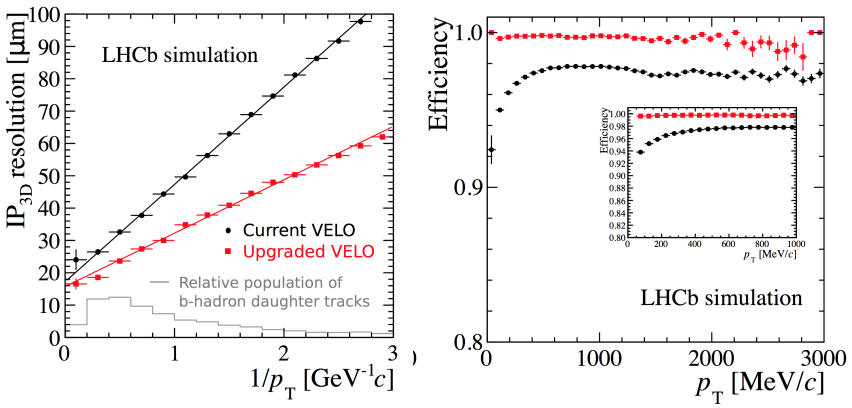


Figure 3.2: Three-dimensional impact parameter resolution as a function of inverse of transverse momentum (left) and efficiency as a function of transverse momentum (right) for tracks reconstructed by the upgraded (red) and current (black) VELO detector [91].

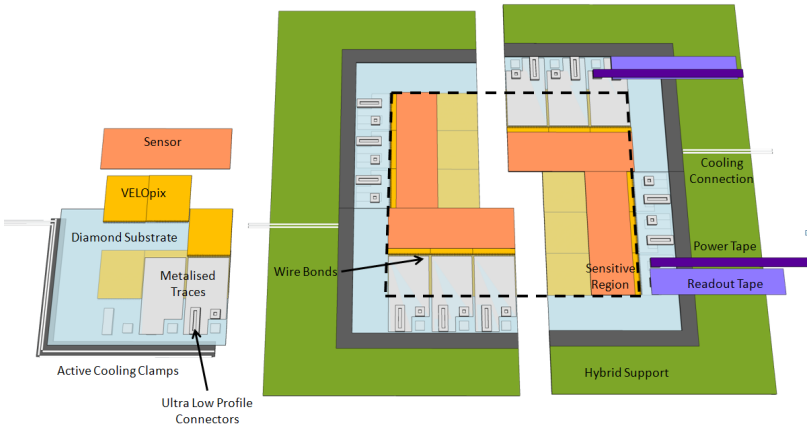


Figure 3.3: Layout of a module of the upgraded VELO detector [92].

and bottom of the detector. The detection mechanism of the SciFi is shown in Fig. 3.4. When a particle traverses the detector, the photons emitted along its trajectory are propagated through the fibres and reach the pixels located at the fibres' end. A signal proportional to the number of fired pixels within a given SiPM channel (each SiPM channel consisting of nearly 100 pixels) is used to determine the position of the particle.

The RICH detectors are currently read out by HPDs that have an embedded 1 MHz readout electronics. For the upgrade, the HPDs will be replaced by commercial multianode PMTs with external readout electronics. In addition, the optics of the RICH1 detector will be optimised to cope with the high occupancy in the innermost region.

The PS and SPD detectors are currently used in the L0 trigger. Given the change in the readout strategy, both detectors will be removed. The front end and back end electronics of the calorimeters will be redesigned according to the new requirements.

The first station of the muon detectors will be removed and additional shielding will be inserted around the beam pipe upstream of the second station. Additionally, the back end electronics of the muon detectors will be replaced since it currenlty provides hit information at 1.1 MHz.

Further details on the upgrade program for the particle identification system can be found in

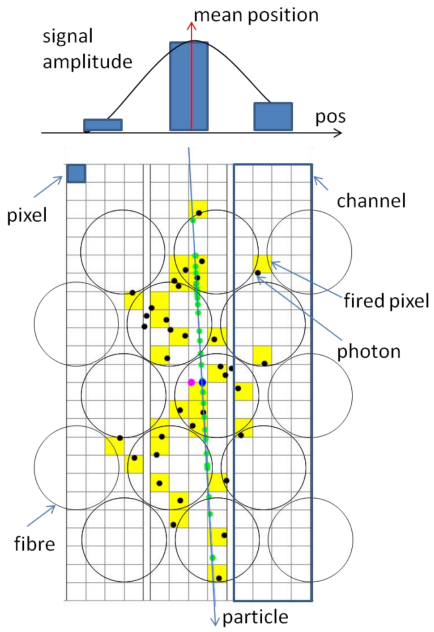


Figure 3.4: Detection mechanism of the SciFi [90]. The circles show the cross section of the scintillating fibres, while the squares show the pixels located at the fibres' end. The rectangle highlighted in blue corresponds to one SiPM channel.

Ref. [93].

The upgraded trigger [94] will consist of two steps: a Low Level Trigger (LLT), which will apply a first selection based on high E_{T} and p_{T}, and a High Level Trigger (HLT), which will apply an offline-quality selection. A real time calibration of the detector will take place at the beginning of each run and the calibrated data will be used to form the trigger decision. The LLT will provide an event-accept rate of 15−30 MHz, while the HLT will guarantee a 20 kHz output rate. An overview of the trigger flow planned for the upgrade is shown in Fig. 3.5.

3.3 The Upstream Tracker

3.3.1 Goals

The Upstream Tracker (UT) [90] will replace the TT [95]. It will play an important role in the HLT tracking, providing a fast estimate of the momentum of charged particles and allowing rejection of low-momentum tracks [96, 97]. Compared to the TT, it will guarantee a higher trigger efficiency, mostly due to its improved acceptance coverage at small polar angles. This will be achieved by reducing the beampipe clearance and insulating material and designing the innermost sensors to have a circular opening around the beampipe, as shown in Fig. 3.6. The UT will also have a finer granularity, improved radiation hardness, and new front-end electronics that will allow a full 40 MHz readout.

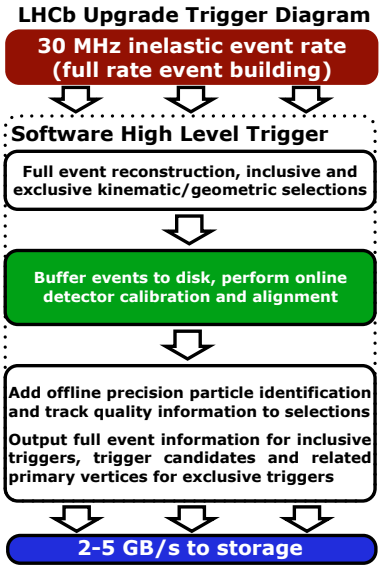


Figure 3.5: Trigger flow planned for the upgrade [84].

3.3.2 Irradiation constraints

The UT is expected to withstand the radiation damage with an integrated luminosity of $50\,\mathrm{fb}^{-1}$ for 10 years of operation. The resulting fluence and radiation dose profiles, obtained by a FLUKA simulation [98, 99], are shown in Fig. 3.7 as a function of the y-coordinate (that is, the vertical coordinate), for a slice of the detector that is positioned at $x = 0$. The x-coordinate defines, together with the y-coordinate, the plane transverse to the beam, with $x = 0$ corresponding to the centre of the beampipe.

The readout chips closest to the beampipe will receive a radiation dose of up to 40 MRad, including a safety factor of 4, while the readout electronics located on the detector frames, at a distance of approximately 70 cm, will be irradiated with up to 100 kRad, including a safety factor of 2.

The innermost sensors will receive a 1-MeV equivalent neutron fluence of up to $3 \times 10^{14}\,\mathrm{n_{eq}/cm^2}$ and will be kept at or below a temperature of about $-5°$C, in order to mitigate the effects of radiation damage and limit the full depletion voltage to below 500 V after irradiation, as demonstrated by studies on prototypes with similar sensors.

3.3.3 Geometry

Similar to the TT, the UT will consist of four detection planes, divided into two sets of two planes each, as is schematically shown in Fig. 3.8 (left). The four planes will be put inside a common light tight box that is flushed with nitrogen or dry air in order to avoid condensation on cold surfaces and that also acts as a Faraday cage.

Each plane will be equipped with single-sided silicon microstrip sensors, having different strip pitches and lengths according to the expected occupancy and read out by 4 or 8 ASICs, depending on the number of strips, as summarised in Table 3.1. Strips will run vertically on the first and last plane and will be tilted by $\pm 5°$ with respect to the vertical direction on the second and third plane.

The functional unit of each plane will be the *stave*, shown in Fig. 3.8 (right) viewed from the

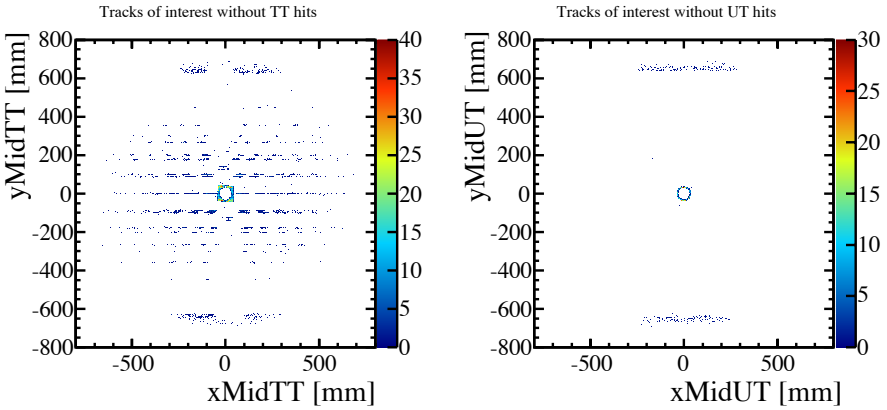


Figure 3.6: Spatial distribution of the charged particles originated from b hadron decays and detected by the VELO and T stations but not by the TT (left) or UT (right). The planes shown in the plots are located in the middle of the TT and UT. The inefficiency of the TT is due to the rectangular opening around the beampipe and to the gaps among adjacent sensors.

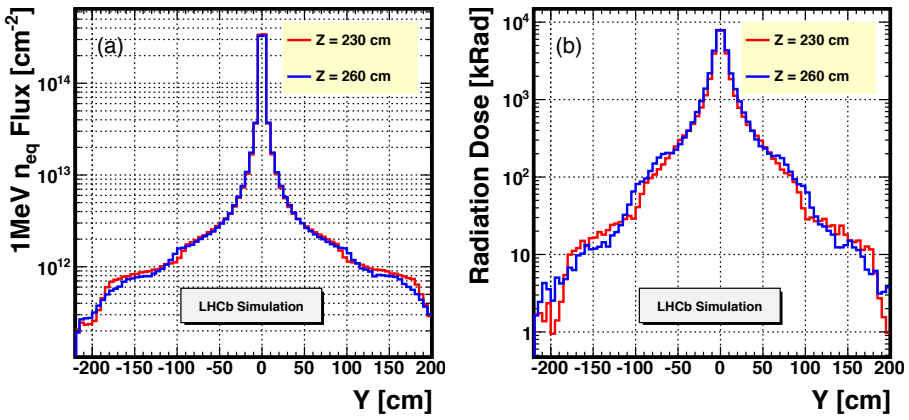


Figure 3.7: Fluence profile (left) and radiation dose profile (right) as a function of y-coordinate, for a slice of the detector that is positioned at $x = 0$ [90]. The red and blue curves refers to the z-coordinates of the two stations of the UT.

front and in Fig. 3.9 viewed from the side. There will be 68 staves in total. Staves within a plane will be staggered in z in order to provide overlap in x and ensure complete acceptance coverage. Each stave will have a width of 10 cm and a length of 134 cm. It will consist of 14 or 16 sensors (depending on its position within the detection plane) mounted on both faces and overlapping in y. Data, control signals, and power will be carried by flex cables running on both faces of the stave from the top and bottom to the centre. Signals will be processed close to the sensor in a newly developed front-end ASIC called SALT [90]. Each ASIC will have 128 channels of front-end amplifiers and will reduce the data volume by performing digitisation, zero suppression, and serialisation output via up to 5 e-ports at a data transmission speed of 320 MBit/s each [100]. Due to the power dissipation in the ASICs, active cooling will be needed inside each stave. A bi-phase CO_2 cooling system with thin cooling pipes embedded in the staves will be used to minimize the impact on the material budget.

The UT will have shorter strips than the TT. Good signal over noise ratio will hence be obtained with thinner sensors, 320 µm instead of 500 µm thick (as in the TT). The increase in the

Type	Pitch (µm)	Nominal length (mm)	Strips	ASICs
Outermost	190	97.28	512	4
Intermediate	95	97.28	1024	8
Innermost	95	48.64	1024	8

Table 3.1: Properties of the different types of sensors.

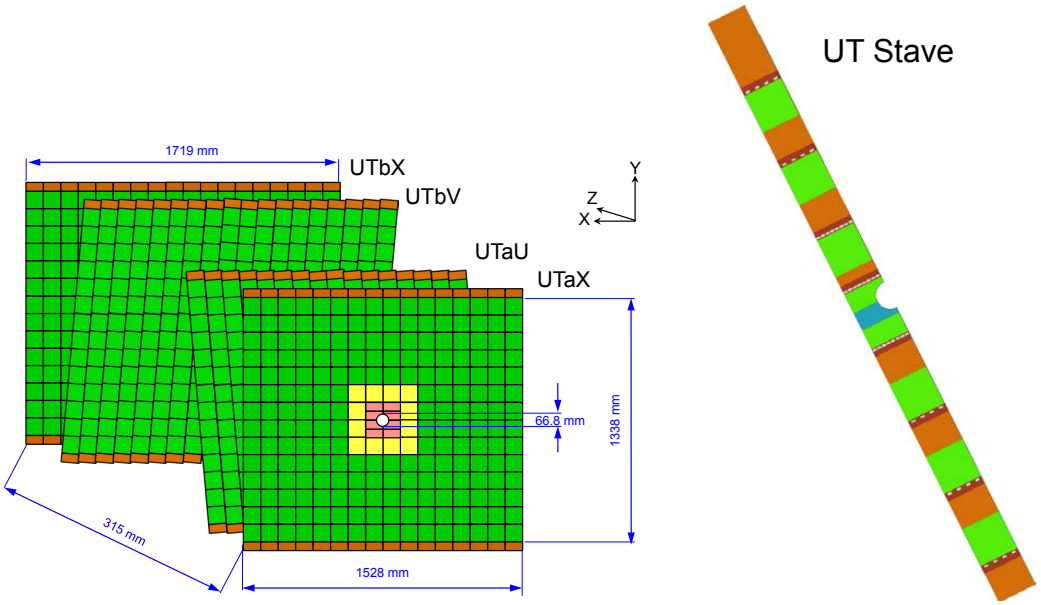


Figure 3.8: Arrangement of the four planes of the UT [90]. Outermost, intermediate, and innermost sensors are shown in green, yellow, and red, respectively (left). Layout of the stave [90]. Sensors are shown in green, flex cables are shown in orange, and ASICs are shown in yellow (right). Sensors, flex cables, and ASICs are also on the other side of the stave.

material budget due to the front-end hybrids and active cooling will therefore be compensated by a corresponding decrease due to the thinner sensors, thus leading to a total material budget comparable to that of the TT.

From the mechanical point of view, the stave will consist of a central support/cooling layer, made of Carbon Fibre Reinforced Polymer (CFRP) facing sheets surrounding a foam core interior in which the cooling pipes are embedded. This central layer of the stave will support the flex cables and detector modules with sensors and ASICs.

3.3.4 Mechanics and cooling

The cooling system will keep the temperature of the sensors below $-5°$C, to suppress reverse annealing, reduce leakage current, and prevent possible thermal runaway due to radiation damage. The maximum temperature difference across the sensor will be kept below 5°C to limit the mechanical stress. The sensors will be kept cold also during shutdown periods, thus minimising reverse annealing. A bi-phase CO_2 cooling system, using thin-walled titanium cooling tubes embedded in the stave, is being designed. The heat load will be mainly due to the ASICs, which is expected to dissipate nearly 500 mW/chip. Therefore, their position is taken into account when designing the path of the

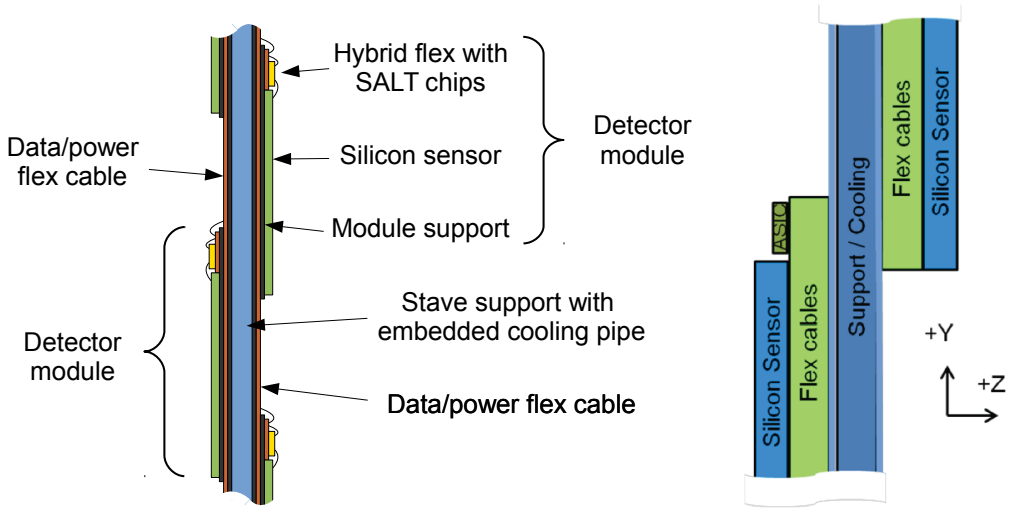


Figure 3.9: Side view of the stave, showing two detector modules, one on each side of the stave (left), and detail of arrangement of support/cooling, flex cables, sensors, and ASICs (right) [90].

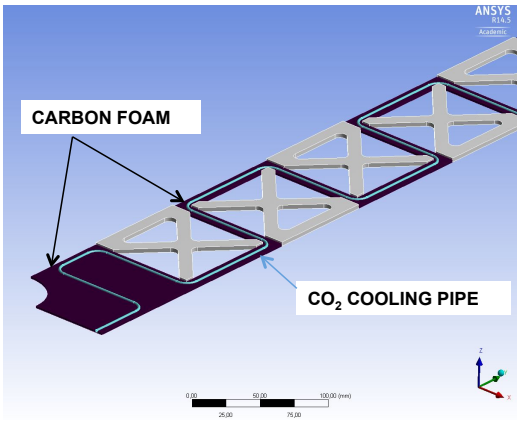


Figure 3.10: Snake pipe design, showing the cooling tubes and the mechanical structure of the stave [90].

cooling tubes. This “snake pipe" design of the cooling tubes, which run directly underneath each row of ASICs, is shown in Fig. 3.10.

3.3.5 Sensors

There will be five different types of sensors, whose properties are summarised in Table 3.1. The innermost sensors will be of two different types, with and without the circular cut out. Two different technologies will be adopted depending on the expected radiation damage: n^+-on-p for sensors in the innermost and intermediate region and p^+-on-n for sensors in the outermost region. However, in order to adopt only one technology within a given stave, the outermost sensors of the central staves are foreseen to be n^+-on-p instead of p^+-on-n. The input capacitance seen by the front-end electronics is expected to be between 3 and 11 pF.

There are three technological challenges involved in the sensor design: the *circular cut out* of the innermost sensors, a *top-side biasing*, and an *embedded pitch adapter*, that is, a second metallisation

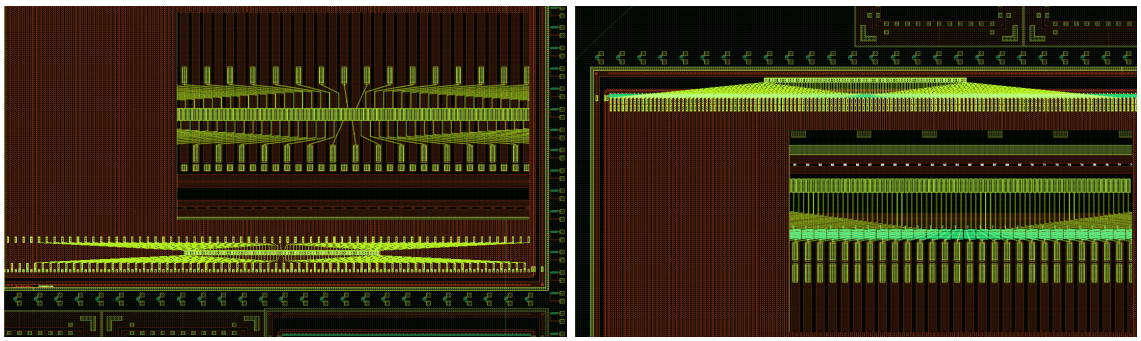


Figure 3.11: Fan-In (left) and Fan-Up (right) geometry of the embedded pitch adapter. A more detailed description of the two geometries is reported in Chapter 5.

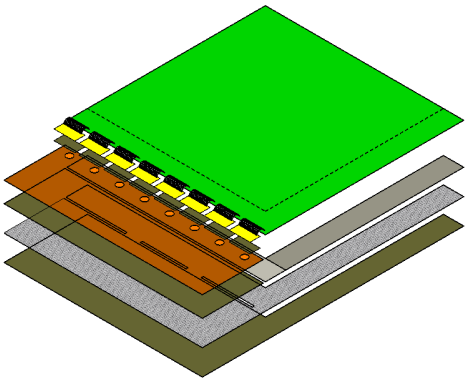


Figure 3.12: Layout of the module. The sensor, in green, and the ASICs, in yellow, are connected through wirebonds, in black. The hybrid flex is shown in orange, the stiffener is shown in shaded grey, and the different types of glue are shown in dark green and grey.

layer, that will allow transitions from the strip pitch of 190 µm on the outermost sensors to the input pad pitch of the ASICs of approximately 80 µm. The embedded pitch adapter presents two main advantages with respect to the conventional solution using an external pitch adapter: reducing the number of wirebonds, thus decreasing the probability of wirebond failures, and reducing the material budget. The top-side biasing and the embedded pitch adapter were tested on prototype sensors, as described in Chapter 5. In particular, two different geometries of the embedded pitch adapter (shown in Fig. 3.11) were tested and resulted in the final design, which is a compromise between the two initial options.

3.3.6 Modules

The design of the modules is shown in Fig. 3.12. A kapton hybrid flex will host the ASICs and will provide electrical connection to the data/power flex cable, to which it will be wirebonded. The ASICs will be glued to the hybrid flex using an electrically conductive epoxy. The hybrid flex will have copper thermal vias to optimize the thermal path from the ASICs to the cooling system. The hybrid flex and the sensor will be glued to a thin support structure, referred to as the *stiffener*, using two different types of glue, differing in their thermal conductivity. This design feature is important to limit the heat transfer from the ASICs to the sensor. The stiffener will extend under the sensor

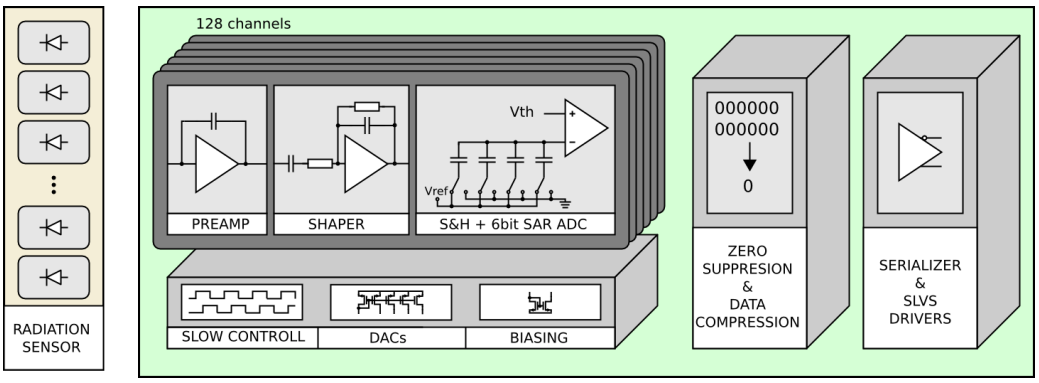


Figure 3.13: Block diagram of the SALT ASIC readout chip [90].

along one of its edges and will consist of Pyrolytic Boron Nitride (PBN), since it has a coefficient of thermal expansion that is very similar to that of the silicon and it is both thermally conductive and electrically insulating. The module will be fixed to the stave using a thin layer of epoxy.

3.3.7 SALT ASIC

The SALT chip, whose block diagram is shown in Fig. 3.13, reads out 128 channels at 40 MHz, is manufactured in radiation-hard TSMC CMOS 130 nm technology, and has an input pad pitch of approximately 80 µm. It has an analogue block, consisting of a preamplifier and a shaper, to guarantee a fast signal, with a peaking time of less than 25 ns and a remainder 25 ns after the peaking time of no more than 5% of the peak value. This design reduces spill over between consecutive LHC bunch crossings. It is optimised for load capacitances of up to 15 pF and has a power consumption of $0.4 - 0.6$ mW/channel. The SALT chip is able to read out both p^+-on-n and n^+-on-p sensors. The shaper is followed by a SAR ADC with 6 bit resolution operating at 40 Ms/s. Once digitized, the data is fed to the digital signal processing block, that allows to mask bad or noisy channels and perform pedestal subtraction, mean common mode subtraction, zero suppression, and data compression. The last block creates and transmits the data frames to the peripheral electronics located on the outer frame of the detector box. Data frames are transmitted through e-links using the SLVS standard at 320 MBit/s data rate. Each ASIC is equipped with 5 e-links, but only some of them are active, depending on the expected hit occupancy on the sensor. The total power consumption of the ASIC is expected to be below 1 W at room temperature.

3.3.8 Flex cable

The flex cable will connect the hybrids to the peripheral electronics on the detector frames, transmitting data, control signals, and power. It will run along the stave, from the top and bottom to the centre, and will have an average length of 0.7 m. The main requirements on the flex cable design are low material budget, low voltage drop, with a maximum round trip drop of 0.5 V, and good signal integrity. The latest prototype design is shown in Fig. 3.14. It consists of three layers of copper traces, with signal traces in the top and bottom layers and power traces in the intermediate layer. All traces are terminated with bond pads that will be connected to the hybrids. On the outer end, the flex cable will be connected to the peripheral electronics through an appropriate connector and pigtail cables.

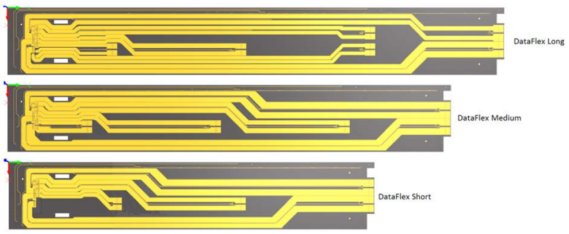


Figure 3.14: Flex cable design. Signal traces are shown in yellow.

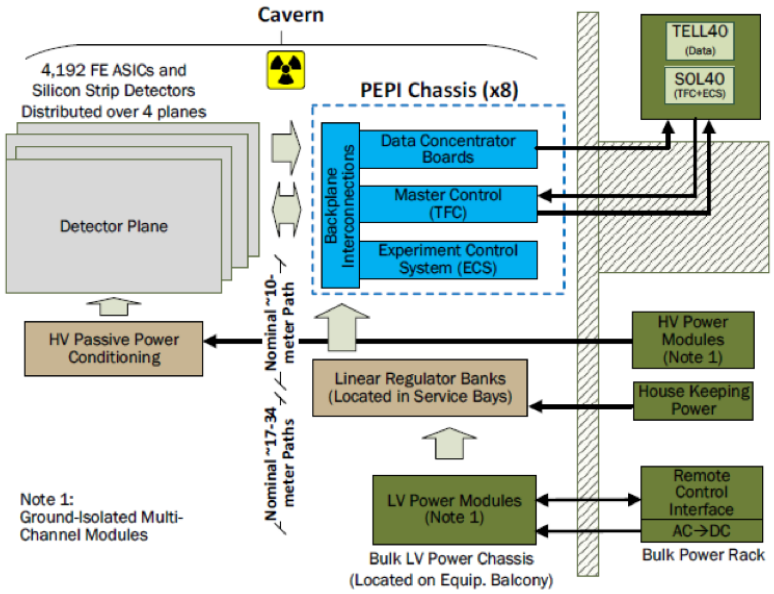


Figure 3.15: Block diagram of the electronics in the cavern and in the counting room [90]. The peripheral electronics are shown in blue.

3.3.9 Peripheral electronics

The peripheral electronics are located at the top and bottom of the detector frames and perform further digital signal processing that cannot be included in the front-end electronics, since it would imply a significant heat load and amount of material in the sensitive area of the detector. They consist of several boards, shown in Fig. 3.15, each with a specific function: electrical to optical transition (Data Concentrator Boards), distribution of fast timing and slow control signals (TFC/ECS), and low/high voltage power conditioning and distribution. After the electrical to optical transition, the data are sent to the TELL40 readout boards located in the counting room, at a distance of 300 m from the detector.

CHAPTER 4

Silicon detectors in particle physics

This chapter describes the working principle of silicon detectors and gives an overview of the main properties relevant to their application in high-energy-physics experiments.

4.1 Working principle of silicon detectors

A silicon detector [101, 102] is a solid-state ionisation counter consisting of a $p-n$ junction to which an external reverse bias voltage is applied. This creates a *depletion region* that is free from charge carriers – electron-hole pairs – and has an electric field directed from the n-doped to the p-doped side of the $p-n$ junction. The thickness of the depletion region d increases with the bias voltage V_{bias} according to

$$d \approx \sqrt{\frac{2\epsilon\epsilon_0}{q}\frac{V_{\text{bias}} + V_{\text{intr}}}{n_{\text{eff}}}},$$

where $\epsilon\epsilon_0$ is the silicon dielectric constant, q is the electron charge, V_{intr} is the intrinsic voltage, and n_{eff} is the effective doping concentration. The intrinsic voltage is the voltage across the depletion region when no bias voltage is applied and is typically about 0.5 V. The effective doping concentration depends on the concentration of donor and acceptor atoms in the silicon detector and has typical values of the order of $10^{12}\,\text{cm}^{-3}$. If the bias voltage is large enough, the depletion region extends to the full sensor and the sensor is said to be fully depleted. The voltage needed for full depletion is called *full depletion voltage* and is approximately

$$V_{\text{fd}} = \frac{q}{2\epsilon\epsilon_0} n_{\text{eff}} D^2,$$

where D is the thickness of the sensor, usually ranging from 200 to 500 µm. Typical values of full depletion voltage range between 100 and 250 V. From the experimental point of view, a good estimate of the full depletion voltage of a sensor can be obtained by measuring its *capacitance* as a function of bias voltage. The capacitance C depends on the area of the sensor A and on the thickness of the depletion region d and is approximately $\epsilon\epsilon_0 A/d$ if the sensor is not fully depleted

and $\epsilon\epsilon_0 A/D$ otherwise. This means that C^{-2} is proportional to V_{bias} if $V_{\text{bias}} < V_{\text{fd}}$ and constant if $V_{\text{bias}} > V_{\text{fd}}$. The full depletion voltage is hence the bias voltage corresponding to the transition between linear and constant behaviour. Typical values of the capacitance are about 1 pF/mm^2 for a 100 µm thick silicon detector. The capacitance described above takes the whole surface of the sensor into account. However, for most microstrip silicon detectors the *interstrip capacitance* is the dominant one, although the exact behaviour depends on the geometry. The interstrip capacitance depends on the ratio of the width and pitch of the implants (described below) and is typically of the order of 1 pF/cm.

Depending on the doping structure, silicon detectors can be divided into p^+-on-n, n^+-on-p, and n^+-on-n. The typical structure of a p^+-on-n silicon detector is shown in Fig. 4.1 and consists of three layers: a thin highly p-doped layer at the top, referred to as p^+, usually segmented in order to provide spatial information, a high-resistivity layer which can be either p- or n-doped in the middle, referred to as the *bulk*, and a thin highly n-doped layer at the bottom, referred to as n^+, which allows an ohmic contact to the backside of the sensor. The n^+-on-p and n^+-on-n silicon detectors consists of three layers too but have a n-doped layer at the top and a p-doped or n-doped bulk, respectively. For p^+-on-n and n^+-on-p silicon detectors, the depletion of the $p-n$ junction starts from the top layer and goes into the bulk, while for the n^+-on-n type it starts from the backplane. The charge carriers collected on the readout strips or pixels are holes for p^+-on-n silicon detectors and electrons for the other types. The main advantages in using n^+-on-p and n^+-on-n silicon detectors are related to radiation hardness and are briefly discussed in the next paragraph.

A *metallisation layer*, usually made of aluminium, is deposited on the silicon surface in order to provide electrical contact to the bias voltage and the readout electronics. Apart from the regions where electrical contact is important, a *passivation layer* of silicon oxide (SiO_2) is used to protect the silicon surface and provide a well-defined interface to the silicon lattice. The elements of the segmented layer at the top of the $p-n$ junction, called *implants*, are capacitively coupled to the readout strips, which are made of aluminium and are located on top of the silicon oxide layer. In addition to the elements described above, silicon detectors have usually other structures, such as bias and guard rings, surrounding the $p-n$ junction. The *bias ring* provides the bias voltage to the implants, usually through a polysilicon resistor. This is characterised by a resistance of about 1 MΩ, which allows to reduce electronic noise and keep all implants at the same potential. The *guard ring*, which is also kept at the same potential of the implants, isolates the $p-n$ junction from the edge of the silicon wafer and guarantees a well-defined electrical boundary for the sensor.

A particle traversing the depleted region produces a number of electron-hole pairs that is proportional to the energy released in the material. The electron-hole pairs are accelerated by the electric field in the bulk and drift toward the surface of the silicon. In the case of a p^+-on-n silicon detector, electrons drift toward the n-doped electrode, that is, toward the backplane, while holes drift toward the p-doped implants. Moving electric charges induce a signal in the implants. Since the latter are capacitively coupled to the readout strips by means of the thin silicon oxide layer, an electric signal is induced in the readout strips and can hence be used to provide spatial information on the particle traversing the detector. The capacitive coupling allows to disentangle the leakage current from the current induced by the incoming particles. For a typical thickness $d = 300\,\mu\text{m}$, a typical electric field $E = 10^3\ V/\,\text{cm}$, and a typical charge carrier mobility $\mu = 10^3\,\text{cm}^2/(V\,\text{s})$ for electrons, the charge collection time is $t = d/(\mu E) \approx 3 \cdot 10^{-8}\,\text{s}$. Mobility is typically $2-3$ times larger for electrons than holes.

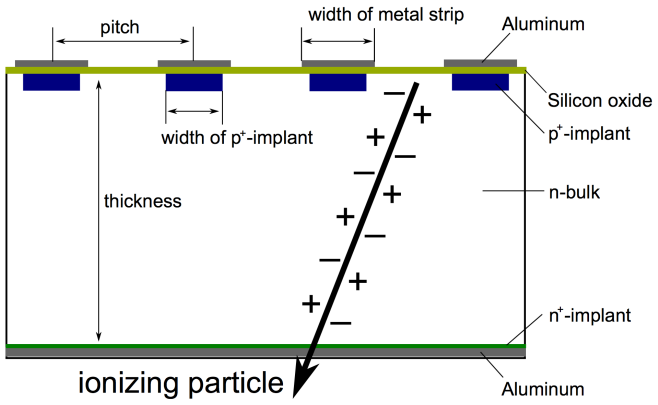


Figure 4.1: Working principle of a silicon detector.

On average, a minimum ionising particle produces 80 electron-hole pairs per 1 µm of path length within silicon. For a typical thickness of 300 µm and assuming that all charge is collected, this results in a signal of 3.84×10^{-15} C. This signal is rather small and is usually processed by a low-noise charge-sensitive amplifier and by a shaper. The amplifier acts as an integrator and converts the narrow current pulse into a step pulse. The shaper, which can be schematically described as the succession of a CR high-pass filter and a RC low-pass filter, limits the pulse width by adding a decay time and increases the rise time by limiting the bandwidth, so that the pulse shape of the output voltage signal has a smooth peak. If the pulse shape is fixed, the output voltage signal has an amplitude that is proportional to the energy lost by the incoming particle and can provide accurate information on the particle's trajectory. The spatial resolution depends mostly on the *pitch*, that is, the distance between two adjacent readout strips. As a first approximation, the dependence of the spatial resolution σ on the pitch p can be written as

$$\sigma = \frac{p}{\sqrt{12}}.$$

Typical values for the pitch are between 40 and 200 µm, which correspond to a spatial resolution from 10 to 60 µm. An improvement of the spatial resolution with respect to these values can be achieved by performing a weighted average of the charge distribution over adjacent readout strips (if charge sharing occurs).

The detector performances described here and others described in detail in Chapter 5 depend on the pitch and width of the readout strips and implants, the properties of the silicon oxide layer, the thickness of the bulk, and the effective doping concentration. In addition, they depend on the readout amplifier and vary over time depending on the radiation damage induced in the silicon detector, as shortly described below.

4.2 Radiation damage

Silicon detectors are affected by two main types of radiation damage: *displacement damage*, due to incident radiation causing a displacement of silicon atoms from their lattice sites, and *ionisation*

damage, due to production of charge carriers and subsequent drift or diffusion toward the silicon oxide interface or other insulating layers, where they can be trapped. Displacement damage, which depends on non-ionising energy loss (NIEL), is considered the main mechanism of radiation damage in silicon detectors, since it alters the properties of the bulk and hence the electric characteristics of the sensors. The displacement damage depends on the type and energy of the incident particles. In order to compare the effects of radiation induced by different types of particles having different values of energy, the radiation damage is usually normalised to a reference particle flux, which consists of neutrons with a kinetic energy of 1 MeV and is referred to as 1 MeV neutron equivalent fluence.

In silicon detectors, the main changes induced by radiation are an increase of the *bulk leakage current*, a change of the full depletion voltage, and a decrease of the *charge collection efficiency* [103].

The leakage current is the residual current across the bulk. It is due to thermally-produced electron-hole pairs and depends on temperature according to

$$\frac{I_1}{I_2} = \left(\frac{T_1}{T_2}\right)^2 \exp\left[\frac{E_g}{2k_B}\left(\frac{1}{T_2} - \frac{1}{T_1}\right)\right], \tag{4.1}$$

where I_1 and I_2 are the leakage currents at temperatures T_1 and T_2, respectively, E_g is the energy gap of silicon (1.12 eV at 300 K), and k_B is the Boltzmann constant. Given the strong dependence on temperature, the leakage current can be reduced by lowering the temperature at which the detector is operated. A factor of 15 is expected between the leakage current at room temperature and the leakage current at $-10°$C. The leakage current depends also on impurities and defects in the crystal lattice, which manifest themselves as mid-gap states, that is, intermediate energy levels between the valence and conduction bands. Displacement damage is due to additional defects originated from radiation. The presence of intermediate energy levels facilitates the transition of electrons from valence to conduction band and causes an increase in the leakage current that can be parametrised as

$$\Delta I = \alpha \Phi_{\text{eq}} V, \tag{4.2}$$

where Φ_{eq} is the equivalent fluence, V is the sensitive volume of the detector, and α is a constant of proportionality that is referred to as *damage parameter*. The damage parameter is obtained by phenomenological models and depends on both temperature and time. This allows to take into account two opposed mechanisms that contribute to further change the leakage current of a detector that has been irradiated: a so-called *beneficial annealing*, which causes a decrease in the leakage current and acts on a time scale of few days at room temperature, and a *reverse annealing*, which has the opposite effect and acts on a longer time scale, of the order of one year at room temperature. In the long term, the reverse annealing has a stronger impact on the detector, so the leakage current increases with time. However, both mechanisms are strongly dependent on temperature and can be reduced by storing and operating the detector at $-10°$C.

The change of the full depletion voltage induced by radiation originates from a change of the effective doping concentration of the silicon bulk. The latter is described by the so-called *Hamburg model*, which is introduced in Ref. [103]. According to the model, the effective doping concentration of an irradiated sensor is the result of a stable damage to the crystal lattice, a beneficial annealing, and a reverse annealing. The stable damage does not depend on temperature and consists of a combination of removal of donors and addition of acceptors, as a consequence of the spatial defects created by the irradiation. For p^+-on-n silicon detectors, effective doping concentration and full

depletion voltage decrease with time, until the silicon bulk becomes p-type and the so-called type inversion is reached. This happens for fluences of nearly $2 \times 10^{13}\,\mathrm{n_{eq}/cm^2}$. After type inversion, the depletion of the bulk starts from the backplane, which means that full depletion of the sensor is required to have charge collection. With time, the silicon bulk becomes more and more p-type, thus causing an increase of the full depletion voltage with time. The full depletion voltage after type inversion eventually reaches the breakdown voltage of the $p-n$ junction, at which point the detector cannot be operated anymore. This limits the lifetime of the detector and is the main reason for relying on n^+-on-p and n^+-on-n silicon detectors when radiation hardness is crucial. The situation, in fact, is different for n^+-on-n silicon detectors, which behave like n^+-on-p after type inversion. Since the depletion of the bulk after type inversion starts from the side of the readout strips, these detectors can guarantee a partial charge collection even without full depletion [104]. This allows to operate the detector with a gradually decreasing efficiency, until the signal-over-noise ratio (described in Chapter 5) drops below the detection capabilities of the readout electronics. No type inversion occurs for n^+-on-p silicon detectors, for which the effective doping concentration keeps increasing with time [105].

The charge collection efficiency, which is the fraction of charge carriers collected on the implants, decreases with increasing levels of radiation due to the trapping of the charge carriers at the defects induced in the crystal lattice. An efficiency loss appears if the shaping time of the readout electronics is shorter than the time it takes to the trapped charge carriers to be re-emitted and then collected by the electrodes. This aspect is investigated in more detail in Chapter 5.

CHAPTER 5

Test beam measurements

This chapter provides a description of the test beam measurements performed on prototype silicon microstrip sensors of the Upstream Tracker. Two main aspects are investigated: the effect of the embedded pitch adapter on the performances of the sensors and the radiation hardness and long-term performances of the top-side and back-side biasing schemes.

5.1 Motivation

As described in Section 3.3.5, most of the UT is made of 320 µm thick, AC coupled sensors with 512 strips and 190 µm pitch. The sensors are coupled to the readout ASICs of 80 µm pitch through an embedded pitch adapter built into the sensors. During fabrication, a second metallisation layer of traces with bond pads matching the readout ASICs pitch is placed on top of the silicon oxide layer, with a direct connection through metal vias to the readout strips below. The readout strips have a width of 123 µm and are separated by a gap of 67 µm. Two different designs of the embedded pitch adapter were investigated between 2014 and 2017. Both types have 64 traces connecting the readout strips to the bond pads of the second metallisation layer. The first type, referred to as Fan-In, is shown in Fig. 5.1. Its traces have a width of 10 µm and a minimum separation of 10−20 µm in the densest region, while the bond pads have a width of 65 µm (perpendicular to the readout strips) and a length of 210 µm (parallel to the readout strips). Both traces and bond pads are located in the active area of the sensor and are separated from the readout strips below by a 2 µm thick insulating silicon oxide layer. The advantage of this design is that the presence of the embedded pitch adapter does not increase the inactive area of the sensor. However, nearly 1 mm of active length along the readout strips has the routing of the embedded pitch adapter on top. This might have an impact on the electrical properties of the silicon detector, especially where the second metallisation layer covers a larger area of the sensor, that is, where the bond pads are located. The second type, referred to as Fan-Up, is shown in Fig. 5.2 and has a similar geometry of the bond pads, but a different routing scheme, with most of the traces and the bond pads located outside the active area of the sensor. Since the second metallisation layer is outside the active area of the sensor, no change in the

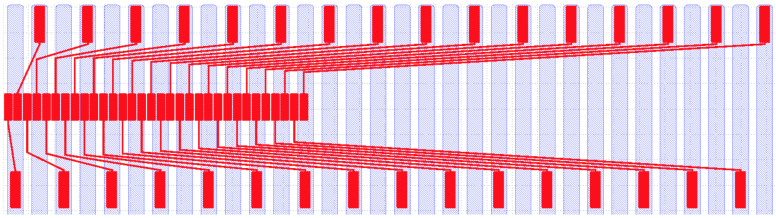


Figure 5.1: Schematic view of one-half of the Fan-In pitch adapter [106]. Blue vertical bands correspond to the readout strips in the first metallisation layer, while red lines and bands correspond to the structure of the second metallisation layer. The bond pads to which the readout ASICs are connected are shown by the red rectangles in the middle. The other one-half of the Fan-In pitch adapter is a mirrored version of this one.

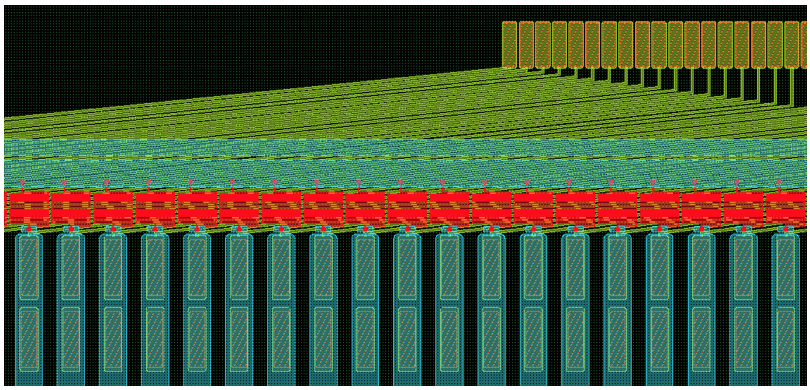


Figure 5.2: Schematic view of one-half of the Fan-Up pitch adapter. Blue vertical bands correspond to the readout strips in the first metallisation layer, red vertical bands correspond to the bias resistors, while green lines and bands correspond to the structure of the second metallisation layer. The bond pads to which the readout ASICs are connected are shown by the yellow bands at the top. The other one-half of the Fan-Up pitch adapter is a mirrored version of this one.

electrical properties of the silicon detector is expected. However, the routing of the pitch adapter adds an inactive length of nearly 500 µm to the sensors.

The biasing scheme of the detector was also investigated. In particular, the possibility to have a top-side biasing instead of the more common back-side biasing was evaluated. In the back-side biasing a part of the backplane is not passivated, so that the high voltage can be applied directly to it by means of a conducting contact. Due to the design of the UT modules, however, routing a conducting contact to the backplane is rather difficult. On the contrary, in the top-side biasing the bias voltage is applied to the top of the sensor. The top-side biasing is achieved by bringing the high voltage to a contact pad located on top of a high voltage ring, which corresponds to the outermost ring shown in Fig. 5.3. The high voltage is then propagated to the backplane through the low resistivity of the peripheral region of the bulk, that is, the region below the high voltage ring. This choice allows to simplify the routing of the high voltage to the sensor. In addition, it allows to completely passivate the backplane, thus decreasing the risk to damage it during handling and preventing accidental short circuits in the high voltage system. However, a degradation in the signal performance could occur with respect to the direct connection to the backplane, due to the resistance of the bulk and the consequent voltage drop, especially after irradiation.

The embedded pitch adapter and the biasing scheme were hence two key elements of the detector design that required dedicated measurements on prototype sensors before production. The first aspect

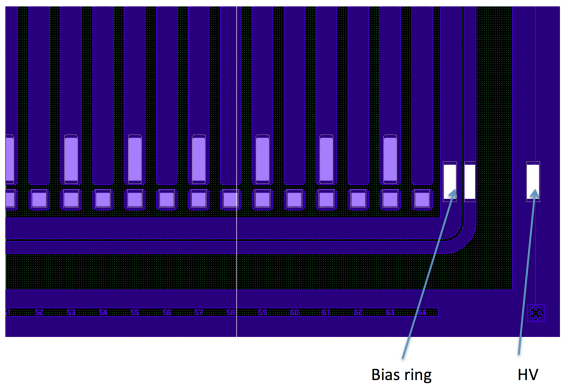


Figure 5.3: Top-side biasing scheme, in which the bias ring connects the strips to ground, while the high-voltage pad, shown on the right, biases the backplane.

that needed to be investigated was the impact of the embedded pitch adapter on the performances of the sensors in terms of charge collection efficiency, cross talk, and radiation hardness. Charge collection efficiency (CCE) is the fraction of charge carriers generated by a particle traversing the depletion region that is successfully collected by the electrodes within the readout time. The embedded pitch adapter modifies the electric field in the region close to the readout strips and, especially for the Fan-In, might result in part of the charge being collected on the second metallisation layer or in a slower signal formation, which would hence cause a loss of part of the signal. Cross talk is the coupling between two or more readout strips through mutual capacitance or inductance. If a particle traversing the detector generates a signal on a given readout strip, an induced signal with a smaller Signal-over-Noise Ratio (SNR) is detected by the neighbouring readout strips too. However, the presence of a double metallisation layer that runs across most of the readout strips of the sensor might induce a non-negligible cross talk not only in the neighbouring readout strips, but also in those placed far away from the signal. In addition, the effect might vary with irradiation. Due to the fact that the embedded pitch adapter occupies only a small region of the active area of the sensor, inefficiency in the charge collection and cross talk are expected to be relevant only for a small fraction of the detected signals. Nevertheless, their impact on the performances of the detector required some specific investigations on both unirradiated and irradiated prototype sensors. The second aspect to be studied in detail was the stability, radiation hardness, and long-term performances of the top-side biasing scheme, compared to the back-side one.

5.2 Silicon sensors

Six mini sensors and one half size sensor produced by Hamamatsu Photonics [107] were tested at the SPS at CERN in May 2016. All sensors are p^+-on-n and have a thickness of 320 µm and a pitch of 190 µm. The mini sensors are 1.4 cm wide and 1.8 cm long and consist of 64 strips each. The half size sensor is approximately 5 cm wide and 10 cm long and consists of 256 strips, of which 106 are wirebonded to the readout electronics described in Section 5.4. The mini sensors have either the Fan-In or the Fan-Up pitch adapter, while the half size sensor has the Fan-Up pitch adapter at one side of the readout strips and the Fan-In pitch adapter at the opposite side, with only the Fan-Up connected to the readout electronics. The backplane of the sensors has no passivation layer in order

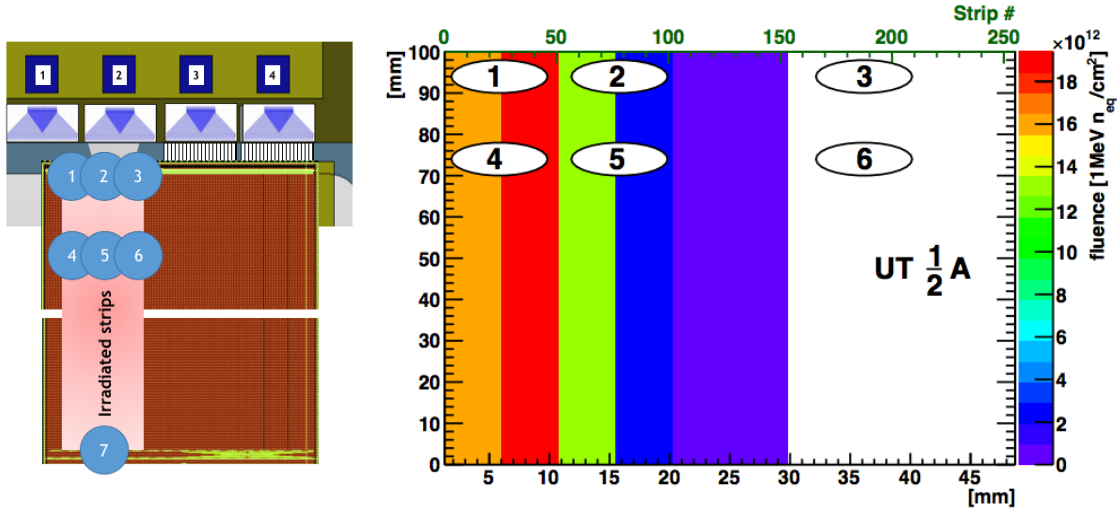


Figure 5.4: Layout of the half size sensor, showing the different regions used in the test beam measurements with respect to the sensor and readout chips (left) and in the local coordinate system of the sensor (right).

to study the performances of the two different biasing schemes.

The measurements were performed at three different irradiation levels, that is, on unirradiated sensors, on sensors irradiated to the nominal fluence of 0.2×10^{14} 1 MeV n_{eq}/cm^2 [90], and on sensors irradiated to twice this value. These values correspond to the expected fluence after $50\,fb^{-1}$ of data taking in the region where these sensors are located, with a safety factor of two, the first, and with a safety factor of four, the second. The sensors were irradiated at the IRRAD facility at CERN using 24 GeV/c protons from the PS. The irradiation profile is Gaussian in two dimensions and has a full width at half maximum (FWHM) of 12 mm ×12 mm. This allows to have a nearly uniform irradiation of the mini sensors, while the irradiation of the half size sensor is performed with the proton beam parallel to the readout strips and is hence uniform along the readout strips but not in the orthogonal direction. Based on this irradiation profile, seven regions were taken into account when characterising the half size sensor, as illustrated in Fig. 5.4. The sketch on the left shows the location of these regions with respect to the sensor and the readout chips, while the plot on the right shows the same regions in the local coordinate system of the sensor, as described in the following. The number of the corresponding readout strips is also shown. Regions 1, 2, and 3 correspond to descreasing fluences in the region of the embedded pitch adapter and were used to study the effects induced by the embedded pitch adapter. Regions 4, 5, and 6 correspond to the same decreasing fluences but in the central region of the sensor and were hence used as control regions. Region 7 corresponds to the intermediate radiation dose on the opposite side of the sensor with respect to the embedded pitch adapter. The latter was used for cross checks and is not discussed further here. The sensors are glued to aluminium carrier boards (shown in Fig. 5.5) that provide support, cooling, and connection to the high voltage system and the readout electronics. The sensors are cooled by two Peltier devices glued to the carrier boards through a thin layer of thermal grease. Condensation was avoided by placing the detectors in a light-tight box flushed with nitrogen. Relative humidity, temperature, and dew point were constantly measured by Resistance Temperature Detection (RTD) sensors. The readout electronics consisted of a TT hybrid hosting 4 Beetle chips, connected to the MAMBA board described in the next sections. A summary of the sensor properties is shown in

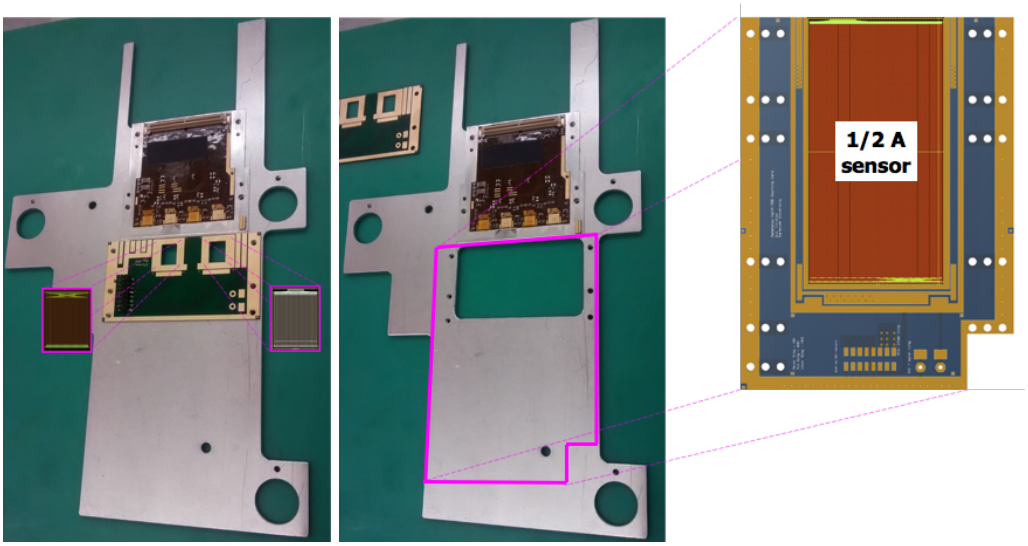


Figure 5.5: Photo of the sensor carrier board for the mini sensors (left) and for the half size sensor (right). The readout electronics is shown at the top and consists of four readout chips, while the location of the sensors is highlighted in pink.

Board	Pitch adapter	Fluence (1 MeV n_{eq}/cm^2)
M1	Fan-In	0.0×10^{14}
M1	Fan-Up	0.0×10^{14}
M3	Fan-In	0.2×10^{14}
M3	Fan-Up	0.2×10^{14}
M4	Fan-In	0.4×10^{14}
M4	Fan-Up	0.4×10^{14}
F1	Fan-Up	0.2×10^{14}

Table 5.1: Properties of the mini sensors, mounted on the M boards, and of the half size sensor, mounted on the F board.

Table 5.1. Additional details on the setup can be found in Ref. [108], which describes analogous measurements performed in the past on n^+-on-p sensors with similar properties.

5.3 Beam properties

The beam consisted of secondary particles produced by the interaction of a high intensity 450 GeV/c primary proton beam with a fixed target made of berillium and lead. The average composition was 67% of protons, 30% of pions, 3% of kaons, and a small fraction of muons. The beam was delivered in spills of nearly 10 s, interspersed with 40 s with no beam. The average energy was 180 GeV and the intensity varied between 10^4 and 10^7 particles per spill.

The spread of the beam was nearly 3 mm along the x axis and 8 mm along the y axis. The beam profile, as reconstructed during a typical data acquisition, is shown in Fig. 5.6. The plots are obtained from the data recorded by the telescope described in Section 5.5. The position of the tracks is evaluated in the local coordinate system of the silicon sensors under test. In this coordinate system, x and y are the direction perpendicular and parallel to the readout strips, respectively. The

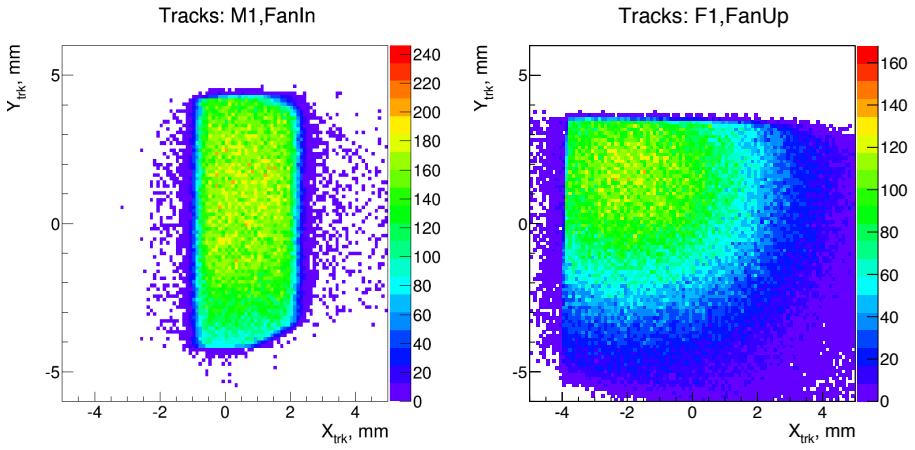


Figure 5.6: Beam profile, as reconstructed by the telescope during the data taking for the M1 sensor (left) and for the F1 sensor (right). Only tracks with a corresponding signal in the silicon sensor under test are shown.

z coordinate corresponds to the position along the beam where the silicon sensors under test are located.

5.4 Readout electronics

The sensors are read out by the Beetle chip, which is the ASIC used in the TT. The Beetle chip processes the signal of up to 128 readout channels using a low-noise charge-sensitive amplifier and a shaper. The resulting analogue pulse has a semi-Gaussian shape that rises to a peak with a peaking time of 25 ns and then falls off with a remainder of the peak voltage after 25 ns of less than 30%. An example of the analogue pulse obtained from laser measurements is shown in Fig. 5.7. The analogue pulse is sampled into an analogue pipeline according to a 40 MHz clock. The analogue pipeline has a depth of 160 samplings, which allows to store the sampled amplitude of the analogue pulse for 4 µs, which is the latency of the L0 trigger of the LHCb detector. A more detailed description of the Beetle chip can be found in Ref. [109]. The sampled amplitude is then processed using the MAMBA board, which has been developed by the INFN in Milan and produced by Nuclear Instruments [110] and is described in detail in Ref. [110, 111]. The MAMBA board digitises the analogue output of the Beetle chip at a fixed time that can be moved in 25 ns increments. The MAMBA board allows to read out one or two TT hybrids, corresponding to up to 8 Beetle chips, for a total of 1024 channels. The readout, which has a maximum rate of 1 MHz, can be triggered by an external signal. In addition, an external clock can be used to timestamp each event, in order to guarantee the offline matching of events recorded by the MAMBA board and by the telescope described below.

The delay between the sampling time of the analogue pulse of the Beetle chip (set by the 40 MHz clock) and the arrival time of the trigger signal formed by the two scintillators described in Section 5.5 can be selected by the DAQ system so that the sampling occurs at the peak of the analogue pulse of the Beetle chip. This is achieved by dividing the 25 ns readout window of the MAMBA board in ten bins of 2.5 ns each, corresponding to different trigger arrival times relative to the 40 MHz clock. The analogue pulse produced by the Beetle chip can hence be sampled at ten different times. This allows to reconstruct the pulse shape inside the 25 ns readout window and to tune the latency of the

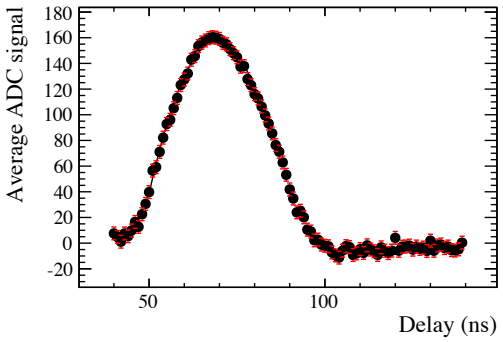


Figure 5.7: Example of the analogue pulse of the Beetle chip, obtained from laser measurements.

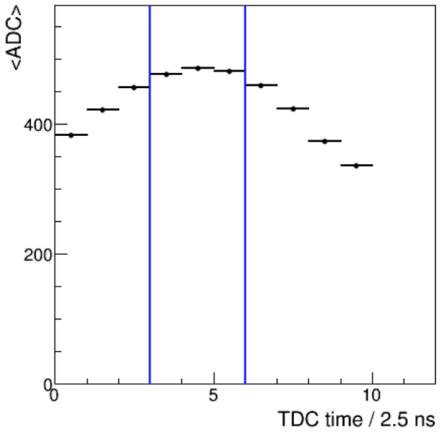


Figure 5.8: Average ADC value of the reconstructed clusters as a function of the TDC time recorded by the MAMBA board, obtained during a calibration data taking for the M1 Fan-In sensor. The vertical lines highlight the peak of the analogue pulse of the Beetle chip.

DAQ system in order to sample the pulse shape at the right time. An example of the calibration is shown in Fig. 5.8. The plot shows the average ADC value of the reconstructed clusters as a function of the TDC (Time to Digital Converter) time recorded by the MAMBA board. The latter is defined as the time shift between the trigger signal and the rising edge of the 40 MHz clock of the Beetle chip. The TDC time is measured in the window between 0 and 25 ns in steps of 2.5 ns.

5.5 Telescope

The Timepix3 telescope [112, 113] has been developed in the context of the LHCb upgrade and subsequently extended to provide a flexible and generic tool for testing the performances of tracking detectors in both high energy and medical physics applications.

Its layout is shown in Fig. 5.9, where the coordinate system is defined such that the z axis corresponds to the beam direction, the x axis is horizontal, and the y axis is vertical. The telescope is made of two separate arms, located downstream and upstream of a Device Under Test (DUT). Each arm has four planes, which are separated by 25 mm and rotated of 9° around the x and y axes in order to improve the single hit resolution. The planes are equipped with a matrix of 256×256 hybrid pixel detector readout chips, the Timepix3 ASICs, bump bonded to p^+-on-n silicon sensors

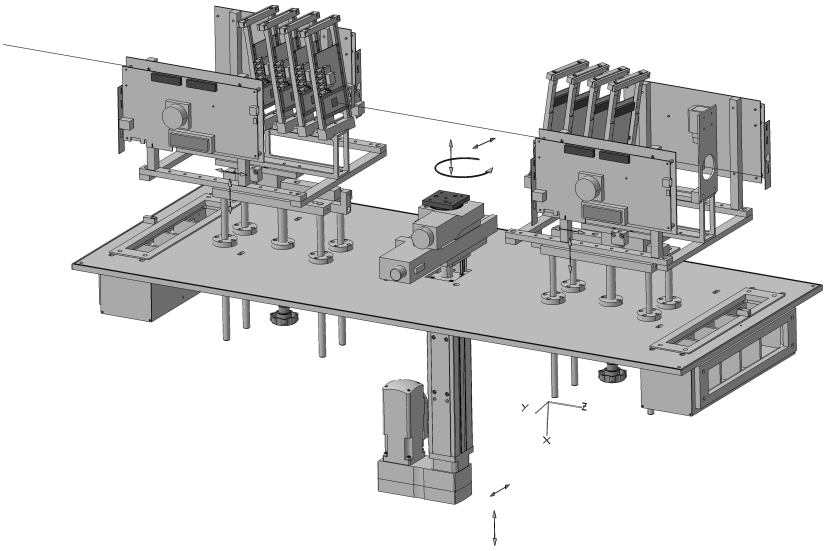


Figure 5.9: Layout of the Timepix3 telescope [108].

of 300 µm thickness and 55 µm pitch along both x and y axes, resulting in a total active area of 14.080 mm × 14.080 mm per plane. Each plane has a material budget in units of radiation length of $x/X_0 = 2.6\%$. The whole telescope can be shifted along the x and y axes, in order to move it out of the beam without the need to physically access the experimental area. Each arm can be shifted along the y and z axes, thus allowing the separation between the two arms to be adapted to the size of the DUT. The latter is mounted on an automatic stage, which allows shifts along the x and y axes with a precision of 1 µm, as well as rotations around the y axis with a precision of 0.01°. In addition, each plane can be shifted independently of the others along the z axis.

Each telescope plane is read out by half a Speedy PIxel Detector Readout (SPIDR) board [114], a readout system developed at NIKHEF, implemented in FPGA and connected to a 10 GBit Ethernet link. The telescope can cope with a rate of 40 million hits / cm^{-2} s^{-1} and provides a pointing precision at the DUT of 1.54 ± 0.11 µm and a timing resolution of 1.56 ns [112, 113].

Two scintillators 1.2 mm high and 1.2 mm wide are positioned upstream and downstream of the telescope and allow to trigger the data acquisition of the DUT. No trigger signal is required to start the data acquisition of the telescope planes, since the Timepix3 ASIC has a data driven readout mode, which sends out a data packet with hit coordinates, Time over Threshold (ToT), and Time of Arrival (ToA). The ToT is related to the charge deposited by the incoming particle, while the ToA allows to match clusters on different planes and hence reconstruct tracks, as well as to associate the reconstructed tracks with the corresponding clusters on the DUT. In order to facilitate the synchronisation between telescope and DUT, the information related to the trigger signal was added to the data recorded by the telescope. In addition, a busy signal was sent from the DUT to the logic forming the trigger signal, in order to inhibit the generation of further trigger signals while the readout electronics connected to the DUT was reading out an event. This ensured that the readout systems of the DUT and of the telescope recorded exactly the same number of trigger signals. Both the Beetle and Timepix3 ASICs have a clock of 40 MHz. The time difference between

these clocks and the trigger signal was measured by a TDC. The TDC time distribution is expected to be uniform for events triggered by incoming particles, since the latter traverse the detector at random times with respect to the clock.

A clustering algorithm was applied to reconstruct clusters out of neighbouring hits on each telescope plane. The cluster position was calculated as the centre of gravity of the hit positions, weighted by the corresponding ToTs. The alignment of the telescope planes and the track fit were performed in two steps, with increasingly stringent requirements. A least squares fit based on Minuit [115] was used in the first iteration, while the Millipede algorithm [116] was used in the second. The least squares fit based on Minuit aligned the telescope planes with respect to the third plane of the upstream arm, allowing translations along the x and y axes and rotations around the z axis. The Millipede algorithm aligned the telescope planes, determining for each of them the corrections corresponding to three translations and three rotations with respect to their nominal position, and simultaneously determined the track parameters. At least one cluster per plane was required to form a track and a cut on the χ^2 per degree of freedom of the reconstructed track was then applied to remove tracks with poor quality. Further details can be found in Ref. [112].

5.6 Measurements

Data acquisitions of up to one million events were taken at bias voltages ranging from 50 V to 450 V for the mini sensors and from 50 V to 210 V for the half size sensor. In the case of the mini sensors, measurements at different angles of rotation around the y axis, in the range between $-10°$ and $20°$, were also taken, while only the nominal angle of $0°$ was considered in the case of the half size sensor. For the mini sensors, the measurements were repeated in two different sectors of the sensor, one with the DUT stage positioned at $y = 0$ or $y = 2\,\text{mm}$, corresponding to the beam illuminating the top edge of the sensor, where the pitch adapter is located, and another with the DUT stage positioned at $y = 10$ or $y = 18\,\text{mm}$, corresponding to the beam illuminating the central region of the sensor. Since the position of the beam changed from one data acquisition to another, two different positions along the y axis were taken into account for each sector. For the half size sensor, seven sectors were considered, as depicted in Fig. 5.4. This allowed to inspect different regions of the pitch adapter and to check if the performances varied as a function of the hit position along the strips. The unirradiated sensors were kept at a temperature between 8°C and 10°C during the measurements, while the irradiated sensors were cooled down to a temperature from -0.5°C to -3.5°C.

The *raw data*, which corresponds to the ADC values measured by each readout channel of the MAMBA board, was processed offline and analysed as described in the next sections. First of all, pedestals and common mode noise, both described below, were evaluated and subtracted from the raw data to obtain the *processed data*. After that, the ADC values measured by adjacent strips were combined together to form *clusters*, corresponding to the position of the particle(s) impinging the detector. Clusters were then used for the high-level analysis, in which some properties of the detector (full depletion voltage, efficiency, charge sharing, cross talk, and resolution) were studied.

Further details on the analysis steps can be found in Ref. [108, 117], which describe similar measurements conducted in 2014 and 2015 on the previous iteration of the UT prototype sensors.

5.6.1 Pedestal calculation

The pedestal is the ADC value measured by a given readout channel in the absence of any signal and noise. This value varies from channel to channel and depends on environmental conditions, like temperature and relative humidity, as well as operating conditions, like the applied bias voltage. For this reason, specific data acquisitions without signal, referred to as *pedestal runs*, were taken right before each measurement, to ensure the proper pedestals were subtracted from the raw data.

The pedestals can be calculated by averaging the ADC values measured by each readout channel during an entire pedestal run. If N is the number of events in a pedestal run and ADC_{ij} is the ADC value measured by the readout channel i in the event j,

$$p_i = \frac{1}{N}\sum_{j=1}^{N} ADC_{ij} \tag{5.1}$$

gives the pedestal that has to be subtracted from the raw ADC value measured by the readout channel i. An improved version of this algorithm consists in first removing some potential outliers that might bias the calculation of the average. This was achieved by fitting the distribution of the ADC values of the readout channel i with a Gaussian function

$$\mathcal{N}_i(x;\mu_i,\sigma_i^2) = N_{0,i}e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} \tag{5.2}$$

and by excluding from Eq. (5.1) all events j for which $|ADC_{ij} - \mu_i| > \alpha\sigma_i$, with α set to three. The standard error of the mean was taken as uncertainty of the pedestal:

$$\Delta p_i = \frac{s_{ADC_i}}{\sqrt{N_{eff}}} = \frac{\sqrt{\sum_{j=1}^{N_{eff}}(ADC_{ij} - p_i)^2}}{N_{eff}}, \tag{5.3}$$

where s_{ADC_i} is the sample standard deviation and the number of effective entries N_{eff} is used.

If not stated otherwise, the pedestal subtracted ADC values are used everywhere.

5.6.2 Noise calculation

Each measurement is affected by a certain level of noise, which originates from both the sensor and the readout electronics. The total noise depends on the properties of the system (for example, the length of the readout strips) and consists of two components, one affecting each readout channel independently of the others and another common to several readout channels.

The average total noise of single readout channels was obtained from the same pedestal runs described in the previous section. In particular, the average total noise of the readout channel i is

$$N_i = \sqrt{\frac{\sum_{j=1}^{N}(ADC_{ij} - p_i)^2}{N}}, \tag{5.4}$$

which has an uncertainty of

$$\Delta N_i = \sqrt{\frac{\sum_{j=1}^{N}(ADC_{ij} - p_i)^2}{2N^2}}, \tag{5.5}$$

corresponding to the standard error of the standard deviation.

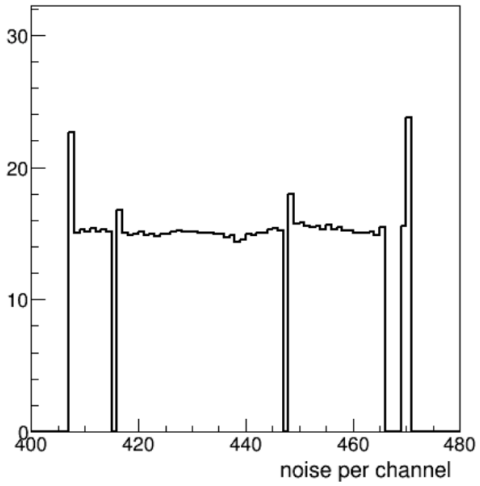


Figure 5.10: Noise as a function of the Beetle channel, obtained for the M1 Fan-In sensor before the CMN subtraction is applied. The vertical axis shows the values of the ADC counts.

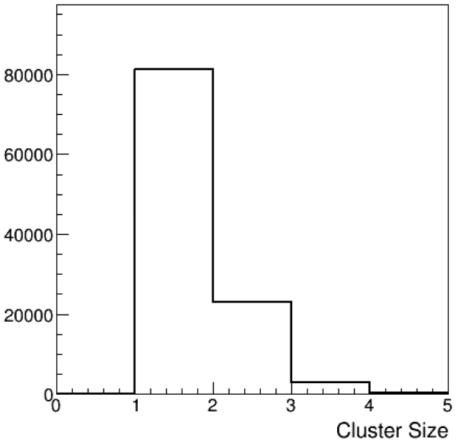


Figure 5.11: Distribution of the cluster size, obtained for the M1 Fan-In sensor above the full depletion voltage. The cluster size of a given bin is shown on the left edge of the bin. The vertical axis shows the number of entries in each bin.

Part of noise affects several neighboring readout channels in a coherent way and is hence called *common mode noise* (CMN). This kind of noise may originate from several causes, including ground loops in the power supplies or readout strips acting like antennas and picking up some environmental noise. It was observed that usually all readout channels of a given chip were affected by the same CMN, although some finer structures were present in some circumstances, depending for example on the number of ports through which the output was transferred out of the readout chip. For this reason, the CMN calculation was performed separately for each Beetle chip. The CMN subtraction is not compulsory when averaging over a large number of events, since the CMN is expected to have a distribution with average zero after pedestal subtraction, but is of crucial importance in all the measurements involving single-event properties, as is the case when calculating the cross talk. A description of how the CMN was calculated and subtracted on an event-by-event basis is reported in the following.

It was observed that the first channels from each readout port of the Beetle chip experience a

Board	Pitch adapter	Connected channels	Masked channels
M1	Fan-In	$407 - 470$	415, 447, $466 - 468$
M1	Fan-Up	$151 - 214$	159, 191
M3	Fan-In	$407 - 470$	–
M3	Fan-Up	$151 - 214$	168, 182, 184, $187 - 190$, 210
M4	Fan-In	$407 - 470$	415, 447
M4	Fan-Up	$151 - 214$	159
F1	Fan-Up	$129 - 234$	–

Table 5.2: Details of connected and masked Beetle channels.

non-negligible cross talk due to the header of the data sent by the chip. These channels, together with those that were not wirebonded to the sensor (and that would hence be affected by a different CMN), were masked in the CMN calculation.

Since the CMN has to be calculated from the same events used in the measurements, two algorithms excluding the signal were developed. One algorithm consists in ignoring the readout channels for which $|ADC_{ij}| > ADC_{th}$, where ADC_{ij} is the ADC value of channel i in event j and ADC_{th} is a fixed threshold that was optimised separately for each sensor. Another way to exclude the signal consisted in ignoring the readout channels for which $|ADC_{ij} - \mu_j| > \alpha\sigma_j$, where μ_j and σ_j are the mean and the standard deviation of the ADC_j distribution (that is, the ADC distribution for the event j) and α is a tunable parameter that specifies the separation between signal and noise. This algorithm is expected to perform better than the other, since a variable threshold allows to deal with both clean and noisy events. Since the mean and standard deviation of the ADC distribution are computed on an event-by-event basis, the resulting threshold is higher for events with a large CMN and lower for events with a small CMN. The performances of the two algorithms were studied on specific data acquisitions and the second algorithm was then chosen for the analysis of the test beam data. In order to further clean the sample from a potential signal, a second iteration was performed.

5.6.3 Masking of noisy or dead channels

The noise as a function of the Beetle channel is shown in Fig. 5.10 for the M1 Fan-In sensor. This distribution is obtained before the CMN subtraction is applied. The gaps correspond to noisy and dead channels, which were masked when analysing the data. The list of connected and masked channels of all sensors is reported in Table 5.2.

5.6.4 Signal-over-noise ratio

The amplifier gain varies from one readout chip to another, which makes it difficult to compare results obtained with different setups. A way to overcome this difficulty is to use the SNR instead of the ADC value. The SNR of a given readout channel i is computed by taking the mean pedestal subtracted ADC value

$$S_i = \frac{1}{N}\sum_{j=1}^{N}(ADC_{ij} - p_i) \tag{5.6}$$

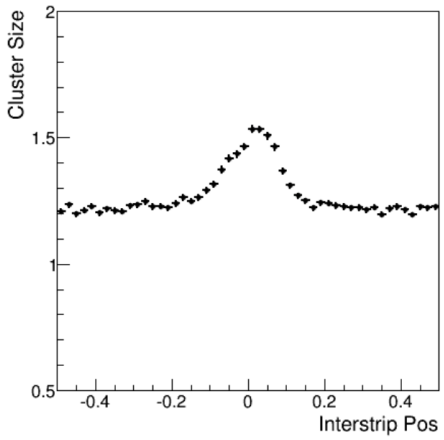


Figure 5.12: Average cluster size as a function of the cluster interstrip position, obtained for the M1 Fan-In sensor above the full depletion voltage. The central values on the y axis correspond to the mean of the distribution of the cluster size in a given bin of the cluster interstrip position, while the error bars correspond to the uncertainty on the mean.

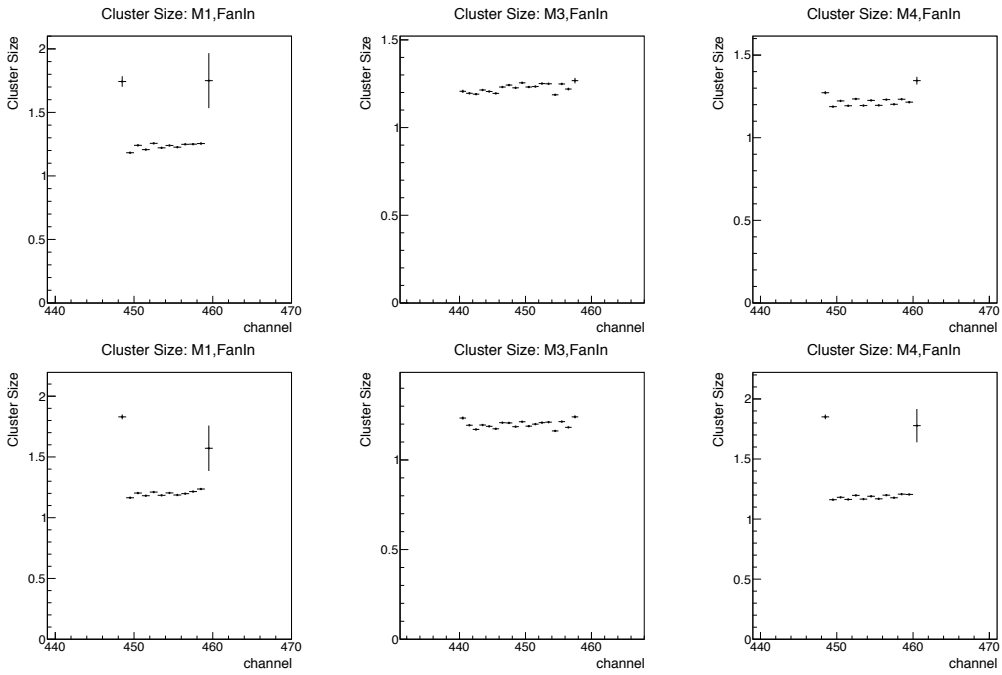


Figure 5.13: Average cluster size of different seed strips of the M1 Fan-In (left), M3 Fan-In (centre), and M4 Fan-In (right) mini sensors, obtained for particles traversing the pitch adapter region (top) or the central region (bottom). The central values on the y axis correspond to the mean of the distribution of the cluster size for a given seed strip, while the error bars correspond to the uncertainty on the mean.

as signal (where the sum is performed over all the events j in a given run) and by dividing it by the raw noise of Eq. (5.4). The uncertainty on the SNR is computed as

$$\Delta(S/N)_i = (S/N)_i \sqrt{\left(\frac{\Delta S_i}{S_i}\right)^2 + \left(\frac{\Delta N_i}{N_i}\right)^2} \tag{5.7}$$

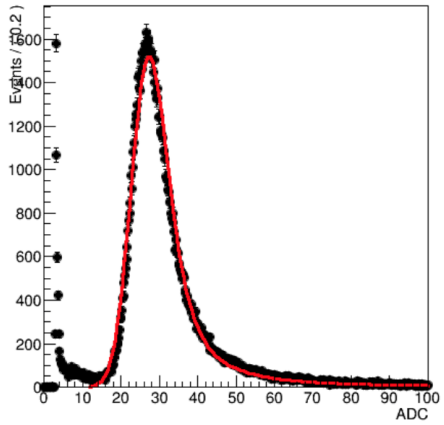


Figure 5.14: Distribution of the ADC values of the reconstructed clusters, obtained during a calibration data taking for the M1 Fan-In sensor, together with a fit of the signal.

since signal and noise can be treated as independent variables. The uncertainty on the signal and the noise is obtained from the standard error of the mean. When considering a cluster of several adjacent readout strips, the sum of the SNR measured by each strip is used:

$$(S/N)_{\text{tot}} = \sum_{i=1}^{n_{strips}} (S/N)_i \tag{5.8}$$

with the uncertainty given by the sum in quadrature of the individual uncertainties.

5.6.5 Clustering

Clusters were built starting from a seed strip with a SNR above 3, to which neighbouring strips were added if their SNR was larger than 2.5. These thresholds are the result of an optimisation procedure based on data with and without signal and depend on the specific properties of the sensors under test. To ensure the results were not biased by the choice of the clustering strategy, several other algorithms were implemented and their performances were compared with those obtained with the default algorithm. The position of the reconstructed clusters was computed as the charge weighted average of the position of the readout strips forming the cluster.

Most of the clusters consisted of one strip only, as can be seen in the distribution of the cluster size shown in Fig. 5.11 for the M1 Fan-In sensor above the full depletion voltage. The cluster size does not depend on the position of the cluster but has a dependence on the *cluster interstrip position*. The latter is defined as the distance of the cluster from the point in between its two nearest readout strips, in units of pitch. According to this definition, the cluster interstrip position has values between -0.5 and 0.5. The maximum cluster size is expected for particles traversing the DUT in between two adjacent readout strips, since the charge sharing is maximised in this situation. The cluster size as a function of the cluster interstrip position is shown in Fig. 5.12 for the M1 Fan-In sensor above the full depletion voltage. A similar behaviour is observed for the other sensors.

The average cluster size might change with the radiation dose. This effect was studied on the mini sensors by comparing the average cluster size of different seed strips for different levels of irradiation. The study was performed for both pitch adapters and no difference was observed between the two.

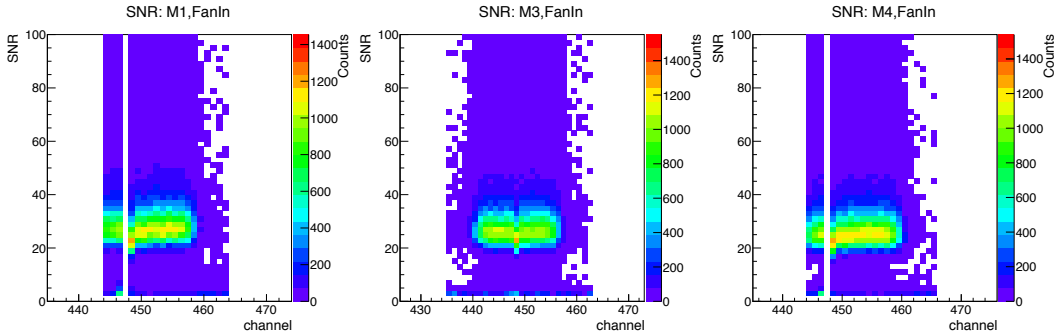


Figure 5.15: SNR of each readout channel of the M1 Fan-In (left), M3 Fan-In (centre), and M4 Fan-In (right) mini sensors obtained for particles traversing the pitch adapter region.

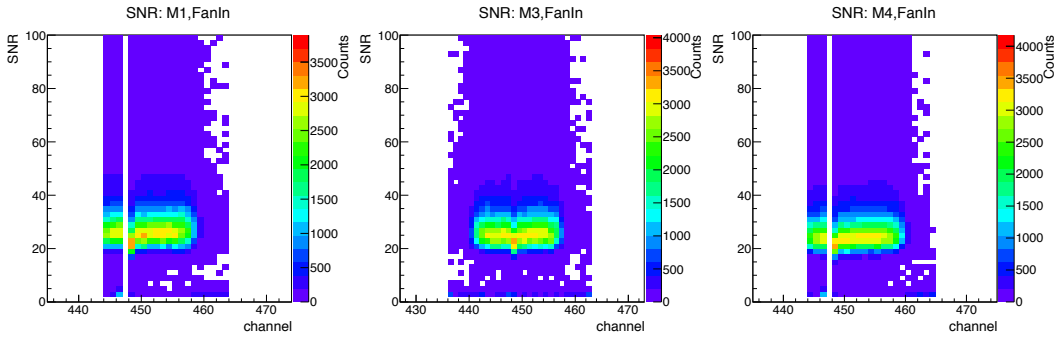


Figure 5.16: SNR of each readout channel of the M1 Fan-In (left), M3 Fan-In (centre), and M4 Fan-In (right) mini sensors obtained for particles traversing the central region.

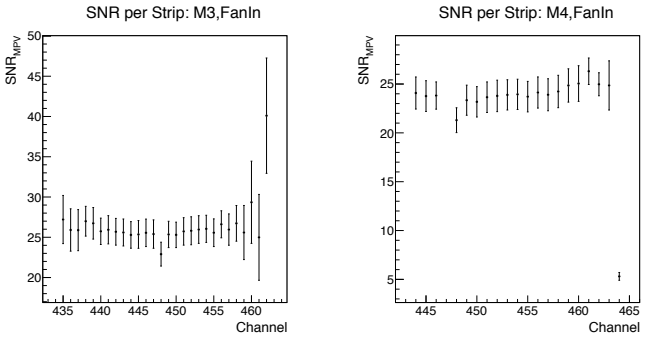


Figure 5.17: Average SNR of each readout channel of the M3 Fan-In (left) and M4 Fan-In (right) mini sensors.

The results are shown in Fig. 5.13 for particles traversing the detector in the pitch adapter region, at the top, and for particels traversing the detector in the central region, at the bottom. The first channel of the M1 and M4 sensors is located next to one of the masked channels and should not be taken into account. The large uncertainties in some bins are due to the low statistics associated to the seed strips that are far away from the beam. The average cluster size is nearly 1.2 for all distributions in Fig. 5.13, that is, it is independent of the position of the incoming particle and does not change significantly with increasing fluences.

An example of the distribution of the ADC values of the reconstructed clusters is shown in Fig. 5.14. This distribution consists of a noise component, characterised by low ADC values, and a signal

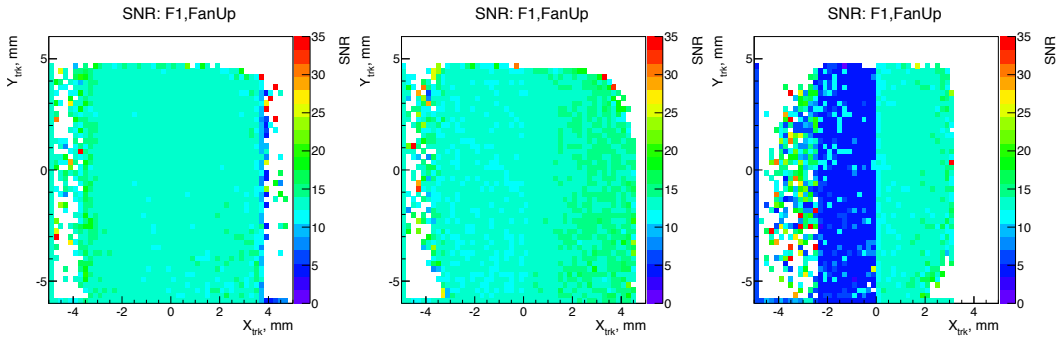


Figure 5.18: Most probable SNR of each readout channel of the half size sensor obtained for particles traversing the regions 4 (left), 5 (centre), and 6 (right).

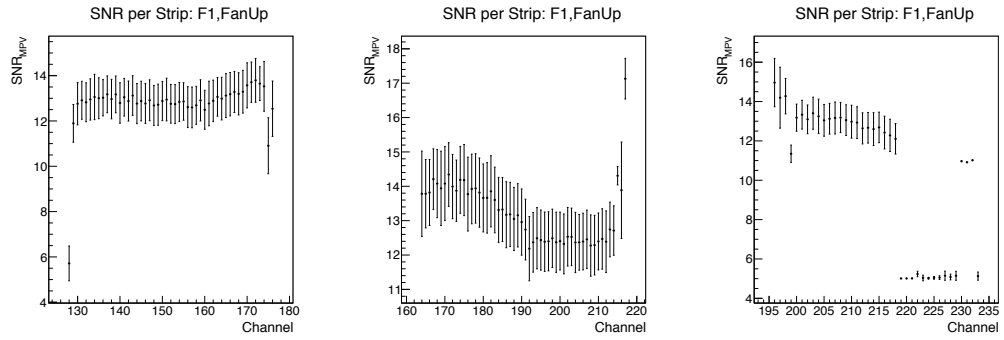


Figure 5.19: Average SNR of each readout channel of the half size sensor obtained for particles traversing the regions 4 (left), 5 (centre), and 6 (right).

component, characterised by a clearly visible peak. The noise component, which peaks at ADC values below 10, corresponds to reconstructed clusters that are not associated to any incoming particle and can be reduced by choosing a more stringent clustering algorithm. However, no attempt was made in this direction since the noise clusters are well separated from the signal clusters. The distribution of Fig. 5.14 was fitted with the convolution of a Landau and a Gaussian distribution. The Landau distribution describes the statistical fluctuations of the ionisation energy loss of a charged particle in a thin layer of matter. The Gaussian distribution describes detector resolution effects. In the fit, all parameters were allowed to float. Namely, these were the mean and width of the Gaussian distribution and the Most Probable Value (MPV) and Full Width at Half Maximum (FWHM) of the Landau distribution. The MPV values obtained from the fit were then used to compute the most probable SNR of the plots shown in Figs. 5.17, 5.18, 5.19, and 5.20.

The distribution of the SNR of each readout channel of the mini sensors with the Fan-In pitch adapter is shown in Figs. 5.15 and 5.16, respectively for particles traversing the pitch adapter region or the central region. White vertical bands correspond to masked readout channels. The profile of these distributions was used to obtain the average value of the SNR of the mini sensors. An example of the profiles is shown in Fig. 5.17 for the M3 and M4 Fan-In sensors. The most probable SNR of the mini sensors was found to be between 25 and 28 before irradiation, between 25 and 27 after irradiation at the nominal dose, and between 23 and 26 after irradiation at twice the nominal dose. No dependence on the position of the beam with respect to the pitch adapter nor on the pitch adapter design was observed. In addition, compatible results were observed when biasing the sensors from the top or from the back. Similar plots for the half size sensor are shown in Fig. 5.18, which

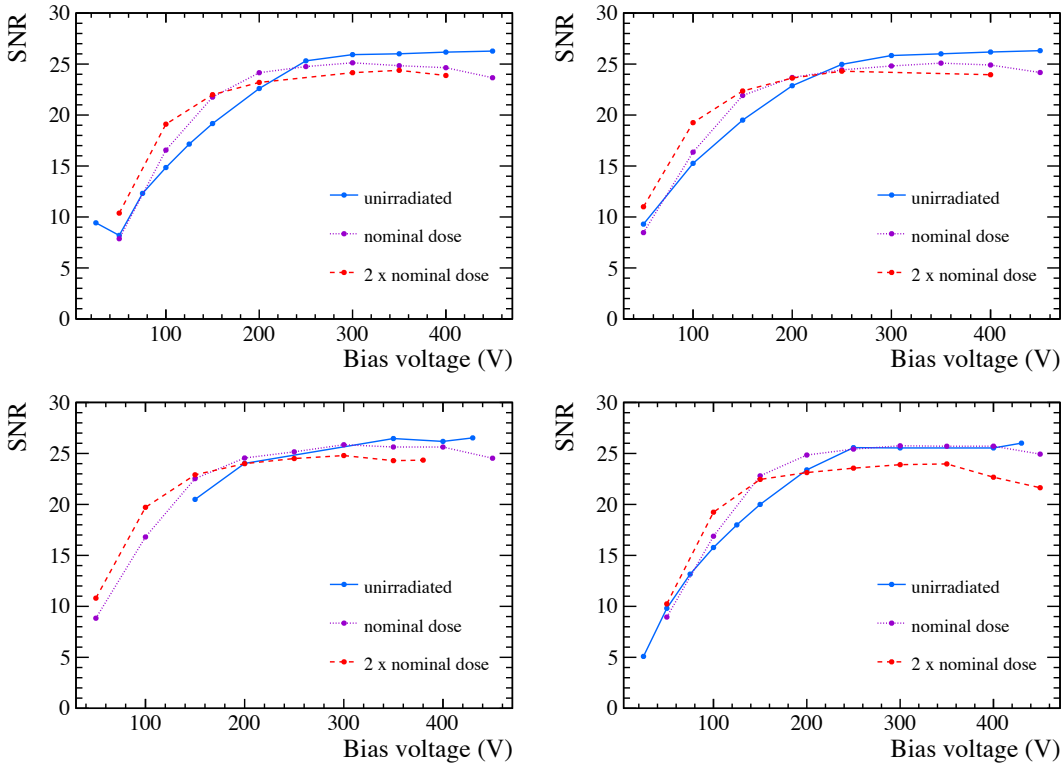


Figure 5.20: SNR as a function of the applied bias voltage for the mini sensors with the Fan-In (top) and Fan-Up (bottom) pitch adapter and with the back-side (left) and top-side (right) biasing scheme. The three curves refer to different radiation doses: no irradiation (solid blue line), nominal dose (dotted violet line), and twice the nominal dose (dashed red line).

displays the map of the most probable SNR as a function of the x and y position of the charged particles. One bin along x corresponds to 200 µm and is hence comparable to the strip pitch of 190 µm. One bin along y corresponds to 250 µm, which is larger than the bond pads of the Fan-In pitch adapter, whose length is 210 µm. Therefore, a possible reduction in the SNR caused by the pitch adapter structures when a particle traverses the pitch adapter region would be diluted due to the other entries in the same bin. The corresponding profiles are shown in Fig. 5.19. It should be noted that the low values in region 6 are due to the fact that the corresponding readout strips are not illuminated by the beam. For the half size sensor the most probable SNR was measured to be between 12 and 15 before irradiation, between 12 and 14 after irradiation at the nominal dose, and between 12 and 13 after irradiation at twice the nominal dose. As for the mini sensors, the results do not show any dependence on whether the beam illuminates the pitch adapter region or the central region. Similar results were obtained for both top-side and back-side biasing schemes.

5.6.6 Full depletion voltage

The bias voltage corresponding to full depletion was determined by studying the most probable value (MPV) of the Landau distribution as a function of the applied bias voltage. In the plots, the SNR values correspond to the MPV of the Landau distribution, as returned by the fit described in Section 5.6.5. The SNR as a function of the applied bias voltage for the mini sensors at different levels of irradiation is shown in Fig. 5.20.

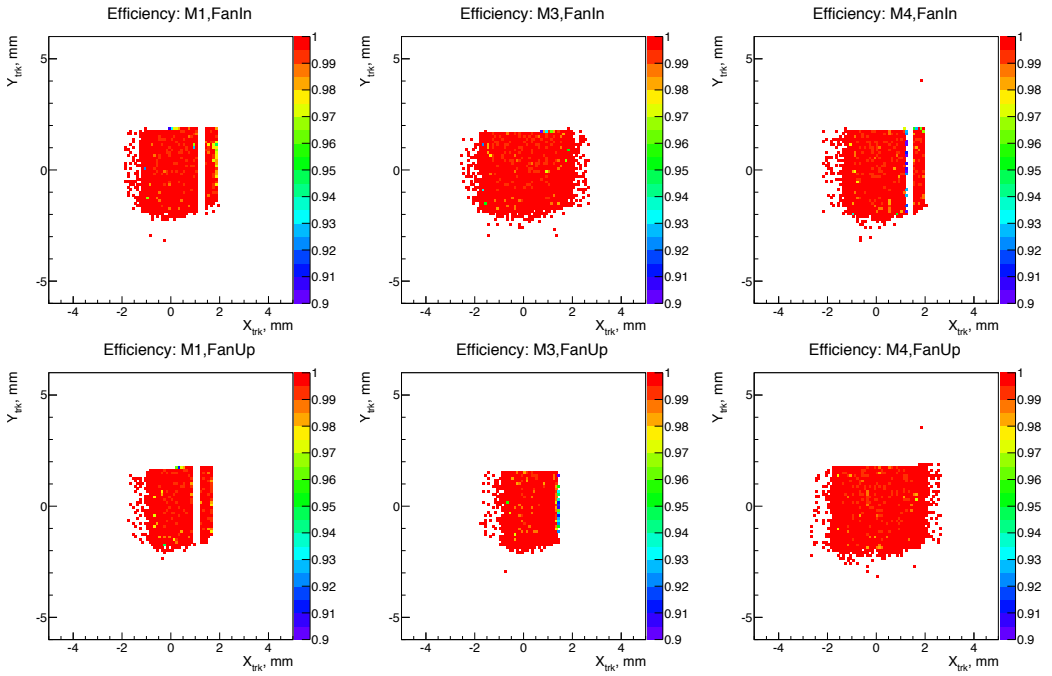


Figure 5.21: Two-dimensional efficiency of the mini sensors in the pitch adapter region, calculated at a bias voltage of 300 V. The distributions refer to the M1 Fan-In (top left), M3 Fan-In (top centre), M4 Fan-In (top right), M1 Fan-Up (bottom left), M3 Fan-Up (bottom centre), and M4 Fan-Up (bottom right) sensors.

For the mini sensors, the full depletion voltage decreases with increasing radiation doses for both pitch adapters and for both biasing schemes. This is the expected behaviour of irradiated p^+-on-n sensors that have undergone a certain radiation damage but have not reached type inversion yet. As expected, no difference was observed for different biasing schemes nor for different pitch adapters.

In order to guarantee a full depletion of the sensors during the measurements, a bias voltage of 300 V was applied to the mini sensors, while a lower bias voltage of 210 V was applied to the half size sensor. At these values of the bias voltage, the leakage current during the data taking was between 30 and 40 µA for the M3 sensor, between 60 and 70 µA or between 230 and 250 µA for the M4 sensor, and between 110 and 120 µA for the F1 sensor. For the M4 sensor, different values of the leakage current were measured depending on the biasing scheme, with the larger values corresponding to the back-side biasing scheme. The same behaviour was observed for both pitch adapters. Since the two M4 sensors are mounted on the same board, these measurements might indicate a problem in the routing of the power lines on the board. In order to understand if the problem appeared with the top-side or back-side biasing scheme, the measured values of the leakage current of the M4 sensors were compared to the expectations based on radiation damage. The expected value of the leakage current was computed by using Eqs. 4.1 and 4.2. The first equation was used to convert the measured leakage current of the M3 sensors (at an average temperature of $-1.55°$C) to the corresponding value at the temperature of the M4 sensors, which was of nearly $-3.15°$C. The second equation was used to obtain the expected increase of the leakage current due to irradiation at the nominal dose and at twice the nominal dose. Based on this calculation, the leakage current at twice the nominal dose is expected to be nearly twice the leakage current at the nominal dose, which is compatible with the values measured with the top-side biasing scheme. The reason of the higher leakage current has not been understood yet.

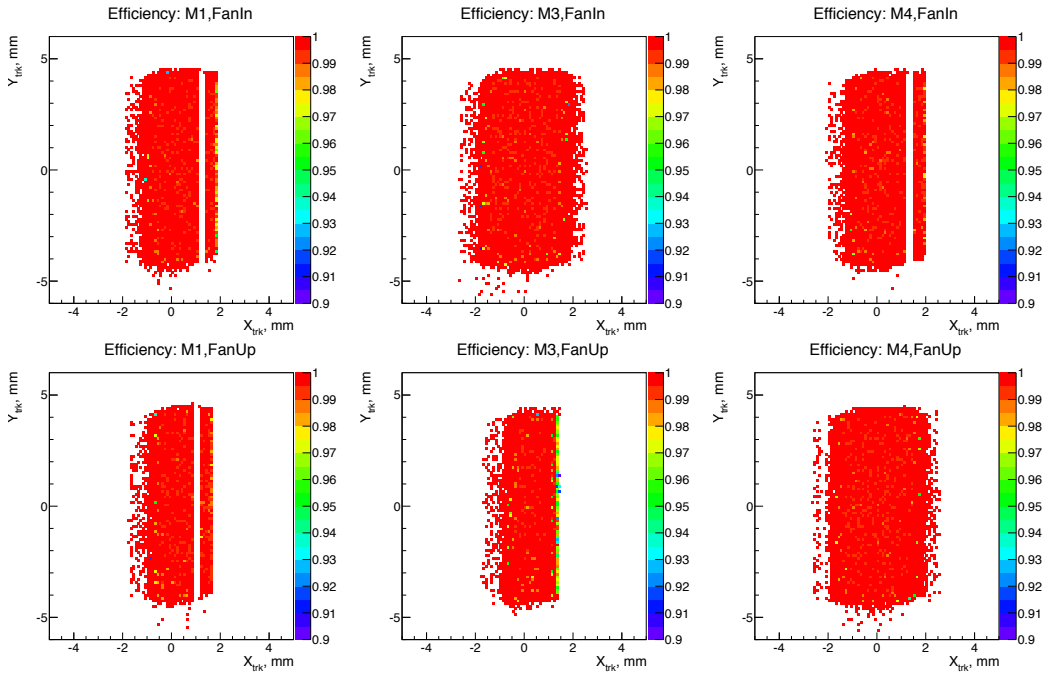


Figure 5.22: Two-dimensional efficiency of the mini sensors in the central region, calculated at a bias voltage of 300 V. The distributions refer to the M1 Fan-In (top left), M3 Fan-In (top centre), M4 Fan-In (top right), M1 Fan-Up (bottom left), M3 Fan-Up (bottom centre), and M4 Fan-Up (bottom right) sensors.

5.6.7 Efficiency above the full depletion voltage

The efficiency was defined as the fraction of tracks reconstructed by the telescope and falling within the DUT acceptance that had an associated cluster reconstructed by the DUT. A cluster was associated to a track if it had a reconstructed x position within 250 µm from the reconstructed x position of the track. The efficiency, which was evaluated above the full depletion voltage of the sensors, was studied as a function of the position of the track on the DUT plane.

The two-dimensional efficiency of the mini sensors, calculated at a bias voltage of 300 V, is shown in Fig. 5.21 for particles traversing the pitch adapter region and in Fig. 5.22 for particles traversing the central region. Increasing fluences are shown from left to right, while the two pitch adapters are shown at the top and at the bottom. The corresponding plot for the half size sensor is shown in Fig. 5.23. The x and y axes show respectively the x and y coordinates of the reconstructed track. One bin corresponds to 100 µm along x and 120 µm along y.

The efficiency of the mini sensors is above 99% for most of the sensor area and no difference based on the fluence nor on the pitch adapter type is observed based on the distributions in Figs. 5.21 and 5.22. The efficiency of the half size sensor slightly decreases with increasing fluences and is lower for particles traversing the pitch adapter region.

The efficiency as a function of the x position of the tracks is shown in Figs. 5.24 and 5.25, respectively for the mini sensors and the half size sensor. For the mini sensors, the efficiency is nearly 99.5% and very stable across the sensor for all irradiation levels, if the particles pass through the central region of the sensor (plots at the top of Fig. 5.24). However, an efficiency loss is visible if the particles pass through the pitch adapter region (plots at the bottom of Fig. 5.24). The efficiency loss is localised in a region of $1 - 2$ mm along x, which is comparable to the width of the bond pads

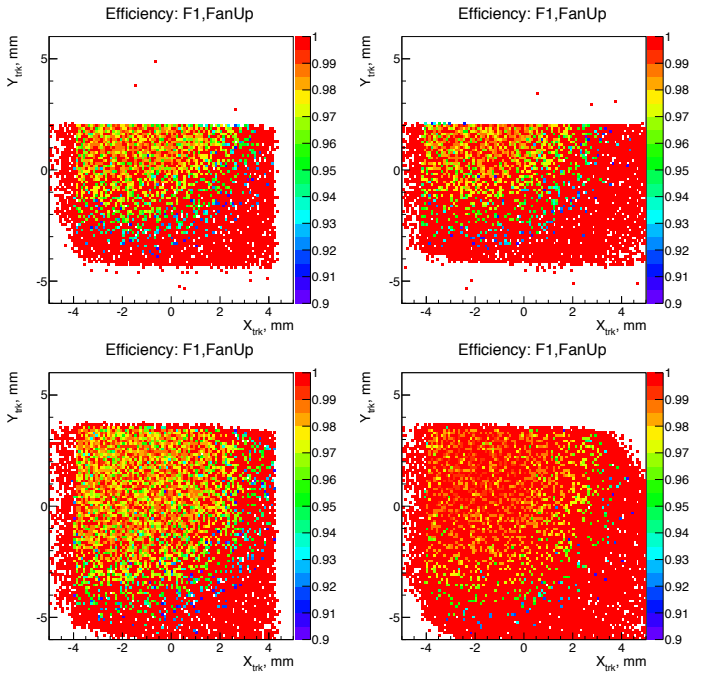


Figure 5.23: Two-dimensional efficiency of the half size sensor, calculated at a bias voltage of 210 V, for charged particles passing through region 1 (top left), 2 (top right), 4 (bottom left), and 5 (bottom right).

of the pitch adapter. The effect is more pronounced for the M3 Fan-In sensor. For the half size sensor, the efficiency is uniform and above 98% for particles traversing the central region of the sensor (plots at the bottom of Fig. 5.25), with a small decrease with increasing fluences. A drop in efficiency of less than 1% is observed when the beam illuminates the pitch adapter region (plots at the top of Fig. 5.25). The effect is reduced, but not totally eliminated, with irradiation.

5.6.8 Charge sharing

In this section, the charge sharing η is defined as

$$\eta = \frac{Q_L - Q_R}{Q_L + Q_R},$$

that is, as the left-right asymmetry of the signals on the two readout strips on the left and on the right of a reconstructed cluster. The signal is given by the SNR. In the following, the readout channel is the one connected to the seed strip.

The η distribution of each readout channel of the half size sensor is shown in Fig. 5.26. The distribution was obtained from a measurement in which the beam illuminated the region 5 of the sensor, which corresponds to the nominal fluence and to the maximum distance from the pitch adapter.

The η distribution integrated over all readout channels of the half size sensor is shown in Fig. 5.27 for particles traversing the regions 4, 5, and 6 of the sensor. These regions correspond to decreasing levels of irradiation and to the maximum distance from the pitch adapter. The charge sharing decreases with increasing fluences and is slightly asymmetric for the unirradiated region of the sensor. The same behaviour was observed for particles traversing the regions 1, 2, and 3 of the

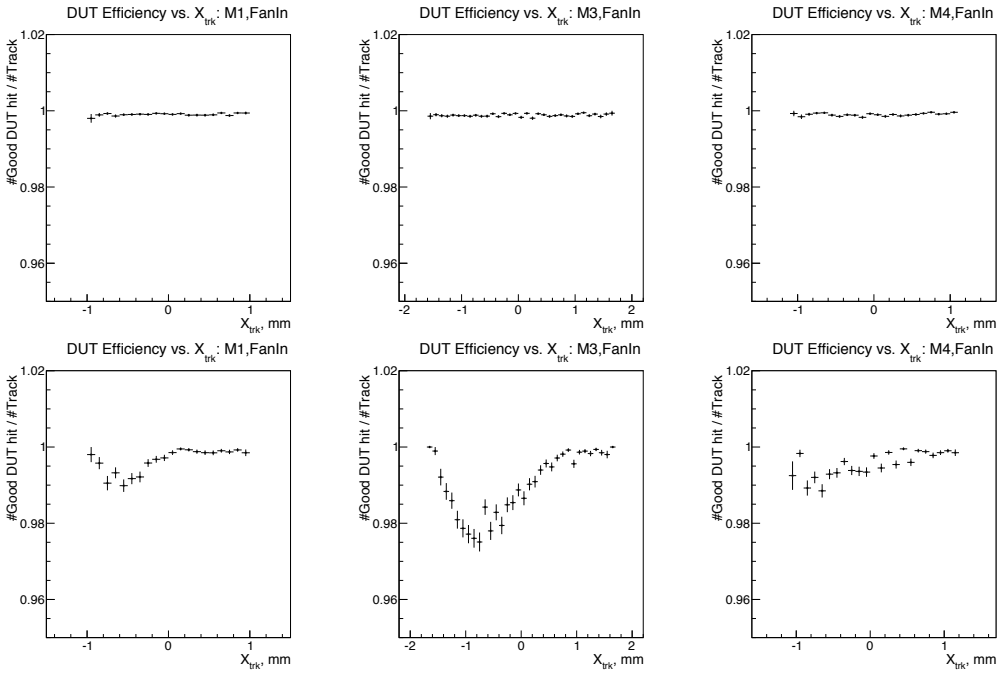


Figure 5.24: Efficiency of the M1 Fan-In (left), M3 Fan-In (centre), and M4 Fan-In (right) sensors as a function of the x position of the tracks, calculated at a bias voltage of 300 V, for charged particles passing through the central region (top) or through the PA region (bottom).

sensor, corresponding to the location of the pitch adapter. The reason behind the asymmetric charge sharing is not understood yet, while the reduction in the charge sharing with increasing radiation doses is consistent with previous observations [118] and might be due to a change in the diffusion properties of the bulk after irradiation. In particular, the mobility of the charge carriers decrease with irradiation. This change can have an impact on the diffusion constant D, which is

$$D = \frac{kT}{e}\mu,$$

where T is the temperature, e is the electron charge, and μ is the mobility of the charge carriers.

5.6.9 Cross talk

In order to quantify the cross talk of the sensors, the following variable has been defined:

$$\alpha(i, \text{seed}) = < \text{SNR}_i/\text{SNR}_{\text{seed}} >,$$

where SNR_i is the SNR associated to the i^{th} readout channel, SNR_{seed} is the SNR associated to the readout channel of the seed strip, and the average is performed over all events having the same seed strip. According to this definition, $\alpha(i, \text{seed})$ is 1 for $i = \text{seed}$ and varies between 0 and 1 for the rest of the readout channels. For a given seed strip, a SNR significantly different from zero is expected on the neighbouring strips that contribute to the same cluster and, possibly, on their neighbours too. A non-negligible value of $\alpha(i, \text{seed})$ in regions far from the seed strip would be an indication of cross talk induced by the pitch adapter.

The two dimensional map of $\alpha(i, \text{seed})$ is shown in Fig. 5.28 for the mini sensors hit by the beam

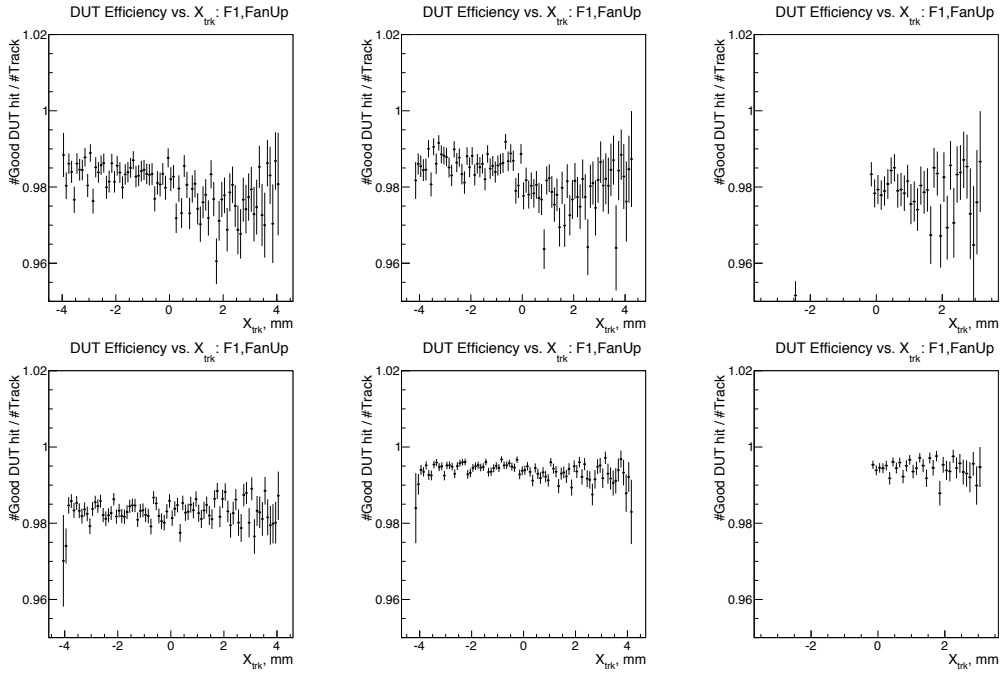


Figure 5.25: Efficiency of the half size sensor as a function of the x position of the tracks, calculated at a bias voltage of 210 V, for charged particles passing through region 1 (top left), 2 (top centre), 3 (top right), 4 (bottom left), 5 (bottom centre), and 6 (bottom right).

in the central region. In the plots, one bin corresponds to one readout channel and the color scale does not go above 0.3 in order to highlight a possible cross talk induced by the pitch adapter. The white bands correspond to the readout channels that were masked during the analysis. It should be noted that only some of the readout channels shown in the plots are associated to readout strips illuminated by the beam. These readout channels are $[444, 458]$ for the M1 Fan-In sensor, $[441, 457]$ for the M3 Fan-In sensor, $[442, 459]$ for the M4 Fan-In sensor, $[188, 198,]$ for the M1 Fan-Up sensor, $[192, 205]$ for the M3 Fan-Up sensor, and $[187, 203]$ for the M4 Fan-Up sensor. The cross talk does not depend on the radiation dose nor on the pitch adapter type. A similar behaviour was observed when the beam illuminated the pitch adapter region. The analogous distributions for the half size sensor are shown in Fig. 5.29 and do not depend on the position of the beam with respect to the pitch adapter. No cross talk above 5% was observed in any sensor.

5.7 Alignment

Several cross checks were performed in order to ensure that the telescope and the DUT were correctly aligned. This is important when matching the reconstructed tracks of the telescope to the reconstructed clusters of the DUT. The procedure to align the telescope is briefly described in Sec. 5.5, while the cross checks to verify the alignment between the telescope and the DUT are discussed below. The results of the alignment studies were taken into account in the measurements described in Section 5.6. However, a detailed description of the alignment procedure requires to know many aspects of the measurements and is hence reported here for simplicity.

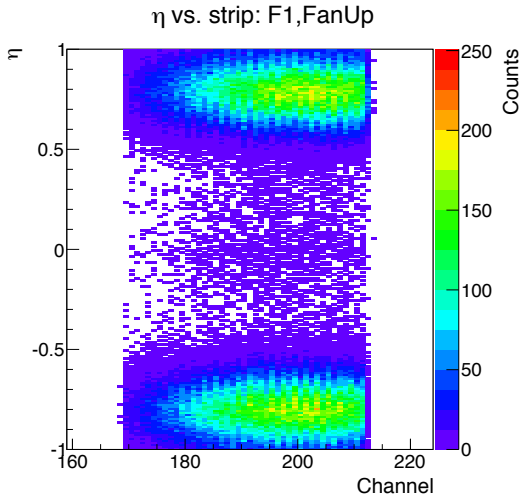


Figure 5.26: Distribution of the charge sharing of each readout channel of the half size sensor, obtained for particles traversing the region 5 of the sensor.

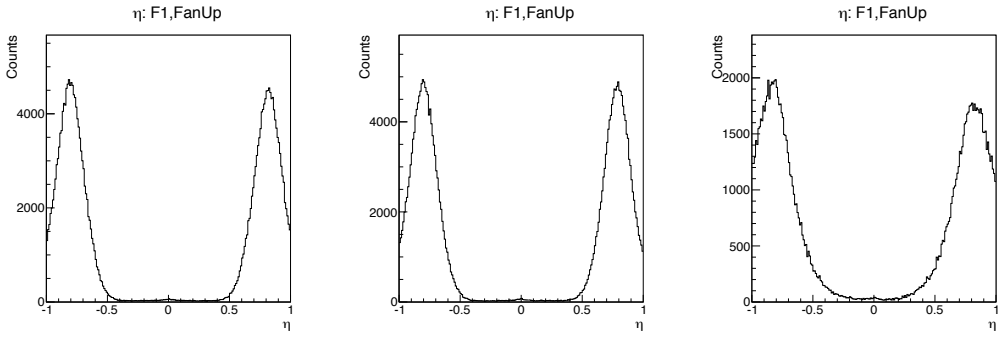


Figure 5.27: Distribution of the charge sharing of the half size sensor, obtained for particles traversing the regions 4 (left), 5 (centre), and 6 (right) of the sensor.

5.7.1 Rotations around the z axis

Potential rotations of the DUT around the z axis were studied by looking at the x residuals as a function of the y position. If the DUT is properly aligned, the x residuals are expected to be independent of the y position, while a rotation of the DUT around the z axis would manifest itself as a linear dependence between the two variables. Two examples, corresponding to a good and a bad alignment, are shown on the left and, respectively, on the right of Fig. 5.30. A straight line fit was performed for each measurement taken at the nominal angle around the y axis. The mean of the angular coefficient distribution was then used to correct for the misalignment when determining the extrapolated position of the track on the DUT plane. The distributions of the intercept a_0 and the angular coefficient a_1, obtained for the M1 board with the Fan-In pitch adapter, are shown as reference in Fig. 5.31. Each entry in the two plots corresponds to a different data acquisition.

5.7.2 Rotations around the y axis

A misalignment of the DUT around the y axis was investigated by measuring the detected signal as a function of the rotation angle around the y axis. The detected signal was defined as the Most

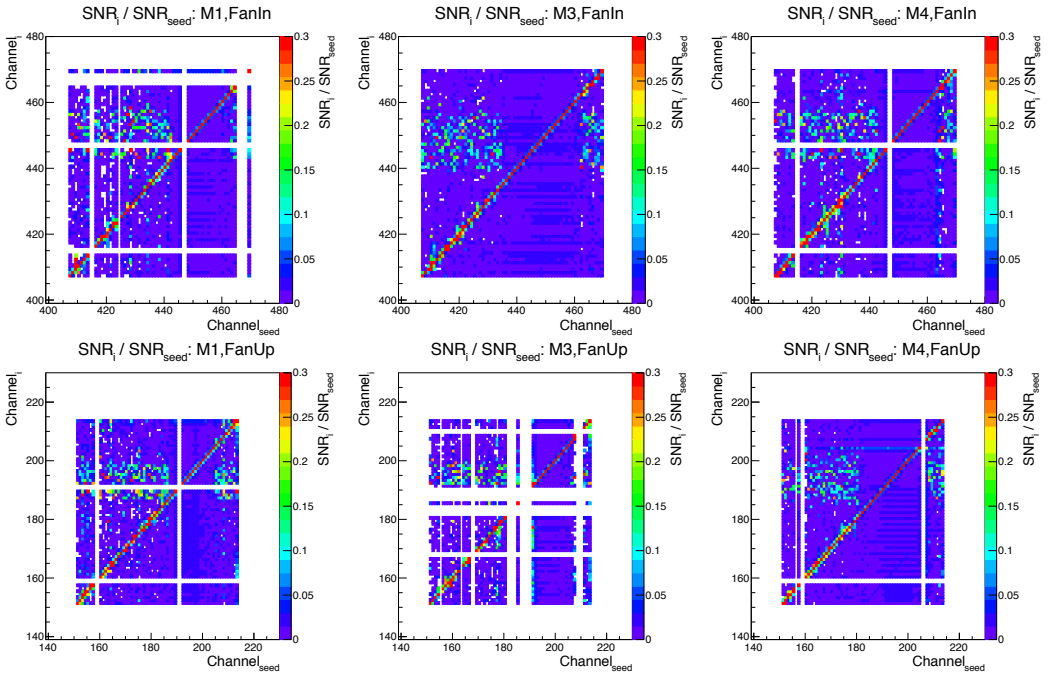


Figure 5.28: Cross talk of the mini sensors for particles traversing the central region, calculated at a bias voltage of 300 V.

Probable Value (MPV) of the Landau distribution in units of SNR. Its dependence on the rotation angle around the y axis is shown in Fig. 5.32 for the M1 Fan-In (left) and M1 Fan-Up (right) sensors. The MPV is expected to be symmetric with respect to the angle for which the beam hits the DUT at normal incidence. As the rotation angle around the y axis increases, the particles traverse more material and consequently release more energy, which leads to a larger deposited charge and hence a larger MPV. The increase in the material is proportional to the $(\cos\theta)^{-1}$, where θ is the angle with respect to the perpendicular to the DUT plane. The dependence between MPV and rotation angle around the y axis is assumed to be

$$\frac{a}{\cos(\theta - \alpha)}, \tag{5.9}$$

where α is the nominal zero angle and a is the corresponding MPV value. The nominal zero angle is zero if the DUT is not rotated around the y axis. The dependence of the detected signal on the rotation angle around the y axis is fitted with the function in Eq. (5.9). The curve corresponding to the fit is overlaid to the experimental points in Fig. 5.32. The nominal zero angle was found to be compatible with zero for all sensors under test.

Studying the dependence of the fraction of one-strip and two-strip clusters on the rotation angle around the y axis gives a similar cross check. Due to the properties of the beam, most of the clusters consisted of one or two strips at normal incidence, with a negligible fraction of clusters with three or more strips. However, as the DUT rotates around the y axis, the fraction of multiple-strip clusters is expected to increase at the expenses of the fraction of single-strip clusters. By studying this effect, it was possible to detect and correct potential misalignments of the DUT around the y axis. Similarly to what discussed above, the distributions of the fractions of one-strip and two-strip clusters are expected to be symmetric with respect to the nominal zero angle, which can be determined by performing a fit with a second-order polynomial function. This is a very simple hypothesis and

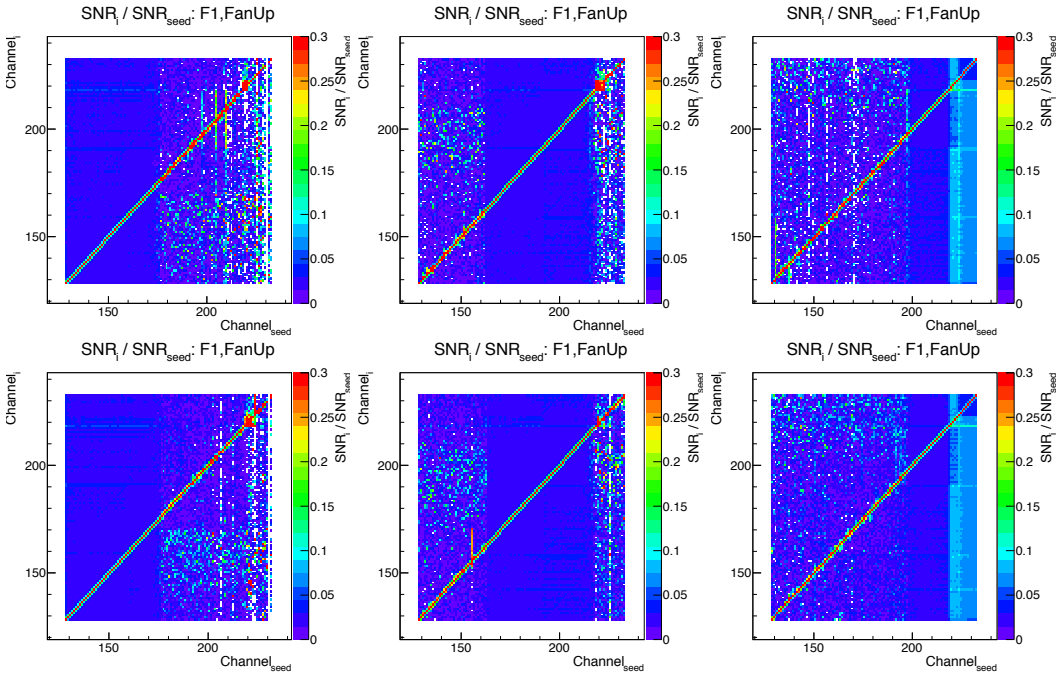


Figure 5.29: Cross talk of the half size sensor, calculated at a bias voltage of 210 V.

does not reproduce the real shape of the distribution for rotations larger than few degrees, but is sufficient to determine the symmetry point. The data and fit results are shown in Fig. 5.33 for the M4 Fan-In (left) and M4 Fan-Up (right) sensors with top-side biasing. Since most of the clusters consist of signal on either one or two strips, the fraction of two-strip clusters has the same angular dependence of the fraction of one-strip clusters apart from a sign.

An alternative method to check the same kind of misalignment consisted in studying the charge sharing as a function of the cluster interstrip position for different rotation angles around the y axis. The rotation around the y axis, that is, the vertical axis, was obtained by rotating the stage where

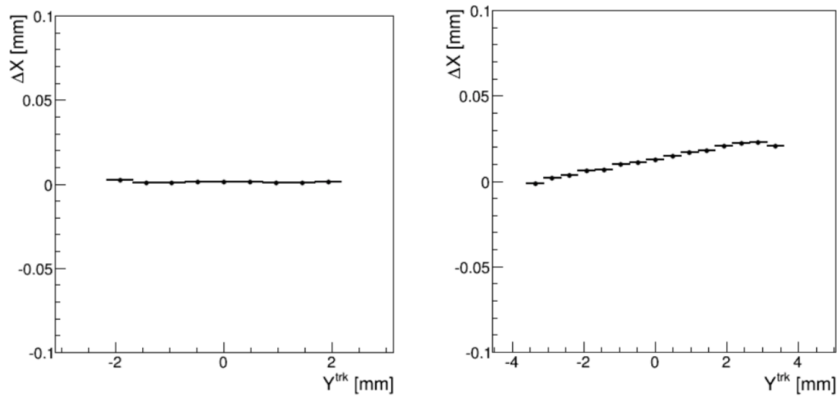


Figure 5.30: Examples of x residuals as a function of the y position for a DUT that is properly aligned (left) and for a DUT that is rotated around the z axis (right).

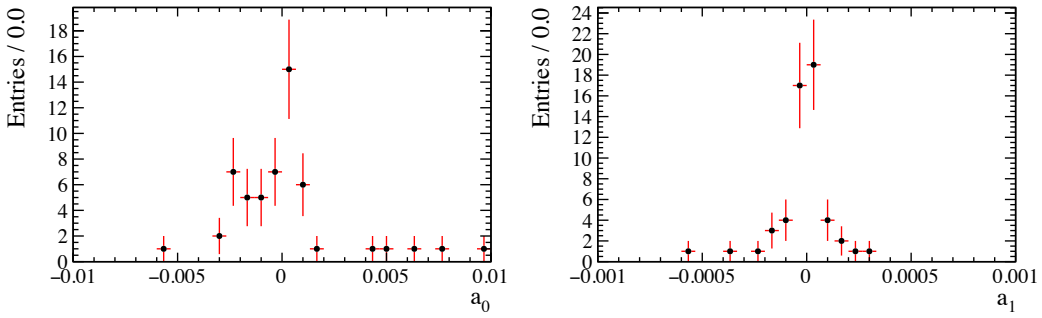


Figure 5.31: Distributions of the intercept a_0 (left) and the angular coefficient a_1 (right), obtained for the M1 board with the Fan-In pitch adapter.

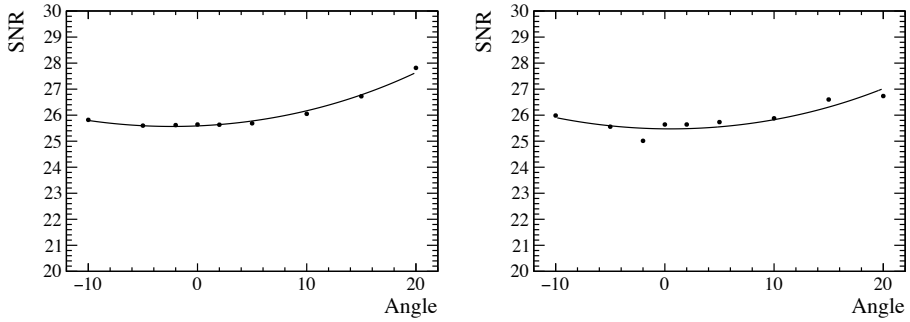


Figure 5.32: SNR as a function of the rotation angle around the y axis for the M1 Fan-In with top-side biasing (left) and for the M1 Fan-Up with back-side biasing (right) sensors.

the DUT was mounted. The charge sharing η is here defined as

$$\eta = \frac{Q_R}{Q_L + Q_R},$$

that is, as the fraction of total charge deposited on the right strip. According to this definition, the charge sharing varies between 0 and 1. The dependence of the charge sharing as a function of the cluster interstrip position on the rotation around the y axis is shown in Fig. 5.34 for the M3 Fan-Up sensor with top-side biasing. Similar plots are available for the other sensors. Each distribution in Fig. 5.34 was fitted with an error function of the form

$$f(x) = a\left(1 - \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right) + b,$$

where a and b are constants, μ corresponds to the mean interstrip position of the falling edge, and σ gives an indication of the width of the falling edge. If the DUT is not rotated around the y axis, the effect of clockwise and anticlockwise rotations around the y axis is the same. The σ as a function of the angle of rotation around the y axis is expected to have a minimum at zero degrees and to be symmetrically distributed around the minimum. The effect of a rotation around the y axis is to increase the charge sharing between neighbouring strips, that is, the value of σ, as the angle of rotation increases. The dependence of the parameter σ on the rotation around the y axis is shown in Fig. 5.35 and is consistent with a symmetric distribution around zero.

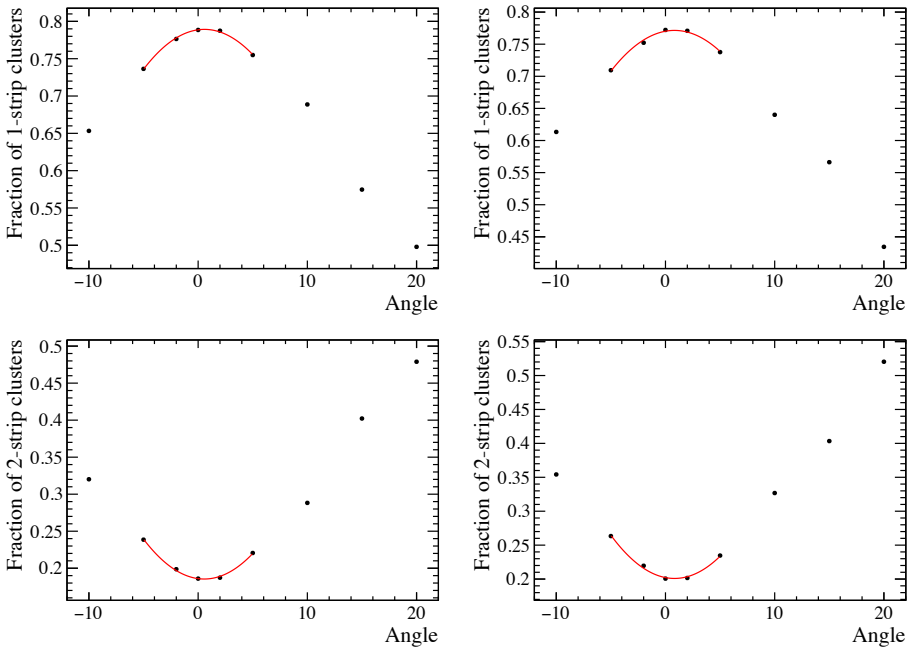


Figure 5.33: Fraction of 1-strip (top) and 2-strip (bottom) clusters as a function of the rotation around the y axis for the M4 fan-in (left) and fan-up (right) sensors, both with the top-side biasing.

5.8 Conclusions

The test beam measurements described in this chapter tested the performances of prototype silicon microstrip sensors for the UT detector. Two main aspects were investigated: the design of the embedded pitch adapter (Fan-In or Fan-Up) and the performances of the top-side biasing scheme.

Seven p^+-on-n prototype sensors with a thickness of 320 µm and a pitch of 190 µm were tested. Six of them are mini sensors and differ in their embedded pitch adapter and fluence. The seventh is a half size sensor with a non-uniform irradiation profile and only one of the two embedded pitch adapters, the Fan-Up.

For the mini sensors, the SNR is measured to be between 25 and 28 before irradiation and decreases to between 25 and 27 at the nominal fluence and to between 23 and 26 at twice the nominal fluence. The SNR of the half size sensor is smaller, between 12 and 15 for the unirradiated regions, between 12 and 14 for the regions irradiated at the nominal fluence, and between 12 and 13 for those irradiated at twice the nominal fluence. The same results were obtained when the beam illuminated the pitch adapter region or the rest of the sensor area. In addition, no dependence on the pitch adapter design was observed and compatible results were obtained when biasing the sensors from the top or from the back. A lower SNR is expected for the half size sensor due to the larger capacitance associated to longer strips.

The full depletion voltage varies from nearly 250 V before irradiation to nearly 200 V and 150 V after irradiation at the nominal fluence and, respectively, at twice the nominal fluence. The decrease of the full depletion voltage with increasing fluences is what one would expect for p^+-on-n sensors before type inversion is reached. The results show that it is possible to operate the sensors for the expected run time of the UT detector.

The charge sharing decreases with increasing radiation fluences and does not show any dependence

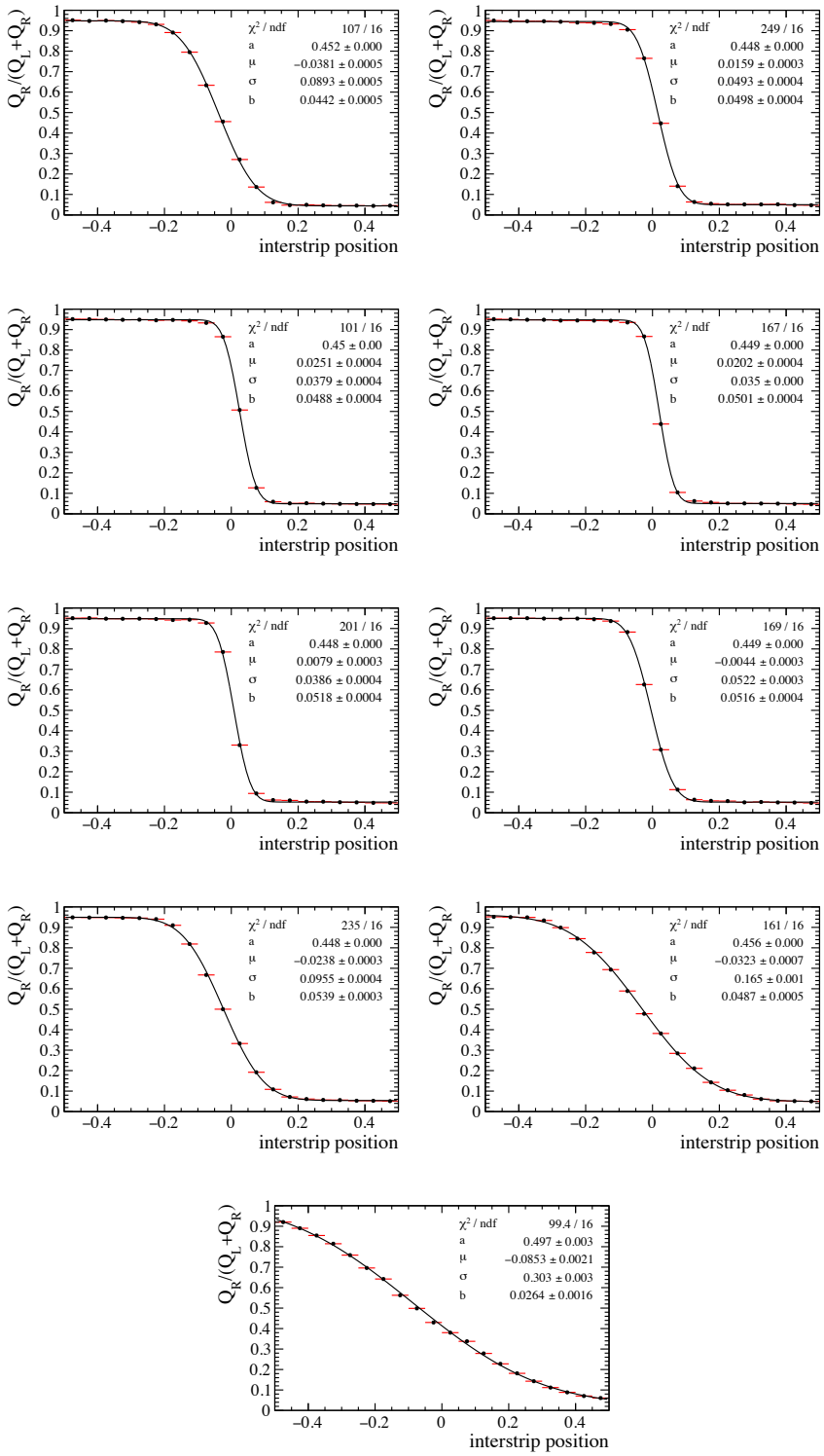


Figure 5.34: Charge sharing as a function of the cluster interstrip position for different rotation angles around the y axis, as obtained for the M3 Fan-Up sensor with top-side biasing. The rotation angles around the y axis are -10° (first row left), -5° (first row right), -2° (second row left), 0° (second row right), 2° (third row left), 5° (third row right), 10° (fourth row left), 15° (fourth row right), and 20° (fifth row).

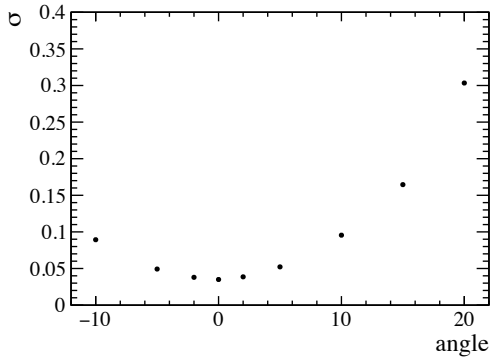


Figure 5.35: Dependence of the parameter σ on the rotation around the y axis, as obtained for the M3 Fan-Up sensor with top-side biasing.

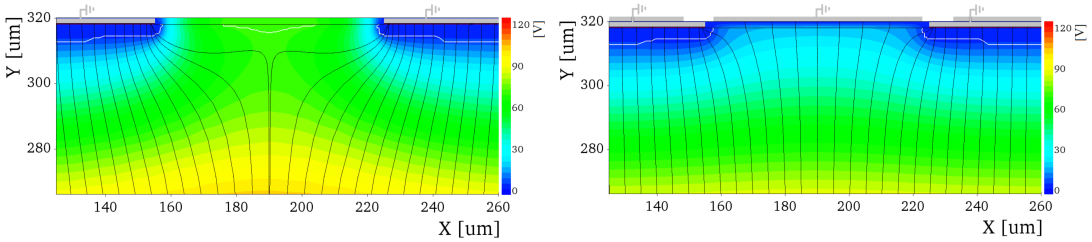


Figure 5.36: Electrostatic potential maps obtained from a simulation of the unirradiated mini sensors tested in the test beam measurements, without (left) and with (right) the Fan-In pitch adapter [106]. The boundary of the depletion region is shown in white. The simulation is based on the TCAD Sentaurus package, which is part of the Synopsys software [119].

on the pitch adapter nor on the biasing scheme.

The cross talk is below 5% for both pitch adapters and does not change significantly after irradiation.

No significant difference was observed between the two biasing schemes in any of the performance tests described above. However, further measurements have shown that the sensors with the Fan-Up pitch adapter have bias voltage stability problems when the top-side biasing is used. The instability consists in a rapid increase of the leakage current when applying a bias voltage between 200 and 300 V. A dedicated simulation was developed to investigate the reason of the instability. It was found that this is related to the high electric field in the transition region between high-voltage ring and bias ring, where the metal traces of the pitch adapter are located. For this reason, a new design was proposed, in which the routing lines are located further from the metal traces of the pitch adapter.

A loss of efficiency localised in the pitch adapter region was observed for the mini sensors with the Fan-In pitch adapter. Since this region is small compared to the active area of the sensor, this effect is negligible (at the per mille level) when integrating over the whole sensor. The localised loss of efficiency does not depend on the biasing scheme and is maximum for the unirradiated sensor. After the test beam measurements discussed here, the effect was studied further on p^+-on-n sensors having the same properties of the mini sensors described in Section 5.2 and was confirmed by dedicated studies on simulation [106]. According to the simulation, the electric field lines in the silicon bulk are modified by the presence of the pitch adapter. This is shown in Fig. 5.36, which compares the electrostatic potential maps with and without the pitch adapter. For simplicity, only two adjacent

readout strips and one bond pad of the pitch adapter were simulated. If the bond pad of the pitch adapter is centred between two adjacent readout strips, the modification of the electric field lines causes a drift of the charge carriers toward the silicon oxide surface, followed by a drift toward the implants. This may result in a loss of signal on the nearest readout strips when a charged particle passes underneath the second metallisation layer. Simultaneously, a signal-like pulse shape may be found on the readout channel connected to the pitch adapter element. As the simulation suggests, the effect has a strong dependence on the relative position between readout strips and bond pads of the pitch adapter. Due to the saturation of the charge density close to the silicon oxide surface, the effect is mitigated by irradiation, although it does not disappear completely [120]. For this reason, the proposed geometry of the Fan-In pitch adapter of the UT sensors has been modified. In particular, the thickness of the silicon oxide layer between readout strips and second metallisation layer has been increased from 2 to 4 µm in order to lower the coupling between the two metal layers.

New prototype sensors with the modified geometry of the Fan-In and Fan-Up pitch adapters have been produced and tested during 2017. No efficiency loss and no bias voltage instability have been observed with the new geometry of the Fan-In and, respectively, Fan-Up pitch adapters. Consistent results have been obtained between the SNR of the prototype sensors described in this chapter and the SNR of the new prototype sensors. However, SNR measurements depend on the readout chip and must hence be repeated using the SALT ASIC, which will be used for the readout of the UT detector. The final geometry of the sensors in the outer region of the UT detector is a compromise between the improved versions of the Fan-In and Fan-Up pitch adapters. The Fan-Up design was preferred, since it allows to keep the embedded pitch adapter outside the active area of the sensor, thus preventing any efficiency loss in the pitch adapter region. The larger inactive area of the sensor, of approximately 500 µm along the readout strips, is compensated by the vertical overlap of the silicon sensors on the two faces of the staves. In the final design, the thickness of the silicon oxide layer between the readout strips and the metal traces of the pitch adapter was increased from 2 to 4 µm, based on the results of the Fan-In pitch adapter.

CHAPTER 6

Analysis strategy and selection of $B^0 \to K^{*0}\ \ell^+\ell^-$ decays

This chapter describes the analysis strategy adopted for the measurement of the ΔS_i observables of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decay, with $\ell = e, \mu$, which is performed at LHCb using the full LHC Run I dataset. The first part of the chapter introduces the counting method approach and its sensitivity to New Physics. The second part discusses the selection of the $B^0 \to K^{*0}\ \ell^+\ell^-$ decays, the techniques adopted to decrease the contamination from several background contributions, the corrections applied to the simulation, and the efficiency calculation.

6.1 Measurement strategy

The measurement of the LFU angular asymmetry observables ΔS_i, Q_i, and D_i is of great interest for the theory community in order to probe NP effects in the $B^0 \to K^{*0}\ \ell^+\ell^-$ decays to electrons and muons.

The measurement of the Q_i observables requires to perform a fit of the *CP*-averaged differential decay width of both $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, as given in Eq. (1.17). While LHCb has already performed this measurement for the muon mode [35], no attempt has been made so far for the electron mode, partly due to the more challenging reconstruction of decays to electrons than to muons. At LHCb, decays with muons in the final state have a clear signature and are reconstructed with high efficiency and good resolution. On the contrary, the reconstruction of electrons is complicated by *bremsstrahlung* emission and other experimental issues which are described in detail in Section 6.2. This results in a lower efficiency and a worse resolution, which limit the possibility to perform a full angular maximum likelihood fit with the statistics collected in the LHC Run I. In addition, the measurement of the Q_i observables requires a good knowledge of the angular dependence of background contributions, since the latter have to be modelled in the fit, and a precise determination of the acceptance as a function of q^2 and the three decay angles θ_ℓ, θ_K, and ϕ.

A simplified approach can be used to measure the ΔS_i observables. This procedure has the advantage of being more robust and less dependent on the statistics available. Therefore, it allows

to measure the ΔS_i observables using the LHC Run I dataset only. This is a first step toward the systematic study of potential NP effects involving LFU violation in the $B^0 \to K^{*0}\,\ell^+\ell^-$ decay and is the main topic of this and the following chapter.

A similar approach can be used to measure the D_i observables, which are however not discussed further in this thesis.

6.1.1 Counting method approach

The $A_{\rm FB}$ and S_i observables, with $i = 4, 5, 7, 8$, can be derived from the symmetry properties of the *CP*-averaged differential decay width of Eq. (1.17). This can be done by considering the following integrations:

$$A_{\rm FB} = \left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_\ell\right) \frac{1}{d(\Gamma+\bar{\Gamma})/dq^2} \frac{d^4(\Gamma+\bar{\Gamma})}{d\vec{\Omega}dq^2},$$

$$S_4 = \frac{\pi}{2}\left(\int_{-\pi/2}^{\pi/2} - \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) d\phi\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_\ell\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{1}{d(\Gamma+\bar{\Gamma})/dq^2} \frac{d^4(\Gamma+\bar{\Gamma})}{d\vec{\Omega}dq^2},$$

$$S_5 = \frac{4}{3}\left(\int_{-\pi/2}^{\pi/2} - \left(\int_{-\pi}^{-\pi/2} + \int_{\pi/2}^{\pi}\right) d\phi\right)\left(\int_0^{\pi} d\theta_\ell\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{1}{d(\Gamma+\bar{\Gamma})/dq^2} \frac{d^4(\Gamma+\bar{\Gamma})}{d\vec{\Omega}dq^2},$$

$$S_7 = \frac{4}{3}\left(\int_0^{\pi} - \int_{-\pi}^{0} d\phi\right)\left(\int_0^{\pi} d\theta_\ell\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{1}{d(\Gamma+\bar{\Gamma})/dq^2} \frac{d^4(\Gamma+\bar{\Gamma})}{d\vec{\Omega}dq^2},$$

$$S_8 = \frac{\pi}{2}\left(\int_0^{\pi} - \int_{-\pi}^{0} d\phi\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_\ell\right)\left(\int_0^{\pi/2} - \int_{\pi/2}^{\pi} d\theta_K\right) \frac{1}{d(\Gamma+\bar{\Gamma})/dq^2} \frac{d^4(\Gamma+\bar{\Gamma})}{d\vec{\Omega}dq^2},$$

where a change of variables is needed, since the *CP*-averaged differential decay width is expressed in terms of $\cos\theta_\ell$, $\cos\theta_K$, and ϕ while the integration is performed with respect to the decay angles θ_ℓ, θ_K, and ϕ. In order to compute the value of a given angular observable, the $B^0 \to K^{*0}\,\ell^+\ell^-$ decays can be divided in two categories, one associated to the positive terms and the other associated to the negative terms of the corresponding formula. The angular observable is then given by the yield asymmetry between the two categories. As an example, the S_5 observable can be computed as

$$S_5 = \frac{4}{3}\frac{S_5^+ - S_5^-}{S_5^+ + S_5^-},$$

where S_5^+ (S_5^-) is the yield of $B^0 \to K^{*0}\,\ell^+\ell^-$ decays in the positive (negative) category. In particular, S_5^+ denotes the yield of $B^0 \to K^{*0}\,\ell^+\ell^-$ decays for which:

- θ_K is in the range $[0, \pi/2]$ and ϕ is in the range $[-\pi/2, \pi/2]$;
- θ_K is in the range $[\pi/2, \pi]$ and ϕ is either in the range $[-\pi, -\pi/2]$ or in the range $[\pi/2, \pi]$.

The complementary subset of $B^0 \to K^{*0}\,\ell^+\ell^-$ decays gives S_5^-. A visualisation of this approach is shown in Fig. 6.1, which represents the distribution of simulated $B^0 \to K^{*0}e^+e^-$ decays in the (ϕ, θ_K) plane, with two different colors highlighting the decays in the two categories. The values of the S_4, S_7, S_8, and $A_{\rm FB}$ observables can be computed in a similar way, according to the following relations:

$$S_4 = \frac{\pi}{2}\frac{S_4^+ - S_4^-}{S_4^+ + S_4^-}, \qquad S_7 = \frac{4}{3}\frac{S_7^+ - S_7^-}{S_7^+ + S_7^-}, \qquad S_8 = \frac{\pi}{2}\frac{S_8^+ - S_8^-}{S_8^+ + S_8^-}, \qquad A_{\rm FB} = \frac{A_{\rm FB}^+ - A_{\rm FB}^-}{A_{\rm FB}^+ + A_{\rm FB}^-}.$$

In particular, S_4^+ denotes the yield of $B^0 \to K^{*0}\,\ell^+\ell^-$ decays for which:

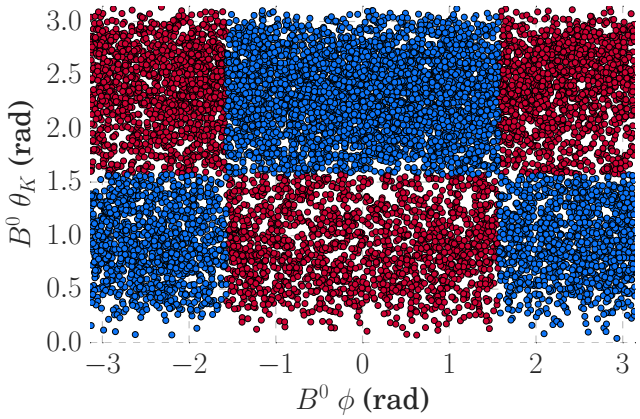


Figure 6.1: Classification of simulated $B^0 \to K^{*0} e^+ e^-$ decays in two categories, S_5^+ and S_5^-, based on the values of the decay angles ϕ and θ_K. The decays in the S_5^+ category are shown in red, those in S_5^- are shown in blue.

- if θ_ℓ is in the range $[0, \pi/2]$, then θ_K is in the range $[0, \pi/2]$ and ϕ is in the range $[-\pi/2, \pi/2]$ or θ_K is in the range $[\pi/2, \pi]$ and ϕ is either in the range $[-\pi, -\pi/2]$ or in the range $[\pi/2, \pi]$;
- if θ_ℓ is in the range $[\pi/2, \pi]$, then θ_K is in the range $[0, \pi/2]$ and ϕ is either in the range $[-\pi, -\pi/2]$ or in the range $[\pi/2, \pi]$ or θ_K is in the range $[\pi/2, \pi]$ and ϕ is in the range $[-\pi/2, \pi/2]$;

S_7^+ denotes the yield of $B^0 \to K^{*0}\ \ell^+\ell^-$ decays for which:

- θ_K is in the range $[0, \pi/2]$ and ϕ is in the range $[0, \pi]$;
- θ_K is in the range $[\pi/2, \pi]$ and ϕ is in the range $[-\pi, 0]$;

S_8^+ denotes the yield of $B^0 \to K^{*0}\ \ell^+\ell^-$ decays for which:

- if θ_ℓ is in the range $[0, \pi/2]$, then θ_K is in the range $[0, \pi/2]$ and ϕ is in the range $[0, \pi]$ or θ_K is in the range $[\pi/2, \pi]$ and ϕ is in the range $[-\pi, 0]$;
- if θ_ℓ is in the range $[\pi/2, \pi]$, then θ_K is in the range $[0, \pi/2]$ and ϕ is in the range $[-\pi, 0]$ or θ_K is in the range $[\pi/2, \pi]$ and ϕ is in the range $[0, \pi]$;

and $A_{\rm FB}^+$ denotes the yield of $B^0 \to K^{*0}\ \ell^+\ell^-$ decays for which:

- θ_ℓ is in the range $[0, \pi/2]$.

One interesting property to note is that the classifications according to S_4 and S_8 can be obtained by looking at those according to $A_{\rm FB}$ and S_5, for the former, and at those according to $A_{\rm FB}$ and S_7, for the latter. In particular, a given decay belongs to the S_4^+ category if it belongs to both $A_{\rm FB}^+$ and S_5^+ or to none of the two, while it belongs to the S_4^- category if it belongs to only one between $A_{\rm FB}^+$ and S_5^+. The same applies to S_8.

In the measurement described in this thesis, the S_i observables are computed as follows:

$$S_i = \frac{S_i^+/\epsilon_i^+ - S_i^-/\epsilon_i^-}{S_i^+/\epsilon_i^+ + S_i^-/\epsilon_i^-}, \tag{6.1}$$

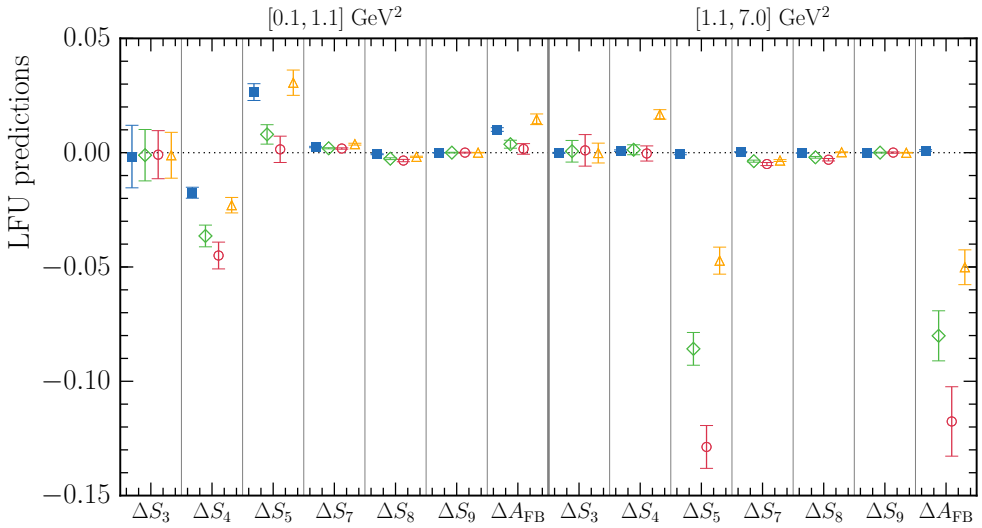


Figure 6.2: Predictions for the $\Delta A_{\rm FB}$ and ΔS_i (with $i = 3, 4, 5, 7, 8, 9$) observables in the low q^2 bin between $0.1\,\text{GeV}^2/c^4$ and $1.1\,\text{GeV}^2/c^4$ (left) and in the central q^2 bin between $1.1\,\text{GeV}^2/c^4$ and $7.0\,\text{GeV}^2/c^4$ (right). The predictions are given for the SM (filled blue squares) and for NP scenarios with $\mathcal{C}_9^\mu = -1$ (in green), $\mathcal{C}_9^\mu = -1.4$ (in red), and $\mathcal{C}_9^\mu = -\mathcal{C}_{10}^\mu = -0.64$ (in orange). The predictions for different models are shifted horizontally for displaying purposes.

where S_i^+ and S_i^- are the signal yields in the S_i^+ and S_i^- categories and ϵ_i^+ and ϵ_i^- are the associated efficiencies. Experimentally, the signal yields in the two categories are obtained from an unbinned extended maximum likelihood fit to the invariant mass distribution of the signal candidates reconstructed in the $K^+\pi^-\ell^+\ell^-$ final state, $m(K^+\pi^-\ell^+\ell^-)$, as described in Section 7.1. The efficiencies, which are not computed on an event-by-event basis but rather as the average over a given region of q^2, are obtained from simulation, as described in Section 6.8. This approach is a valid approximation if the efficiency does not have a strong dependence on q^2. It is important to emphasise here that, since only asymmetries are computed in this approach, the result is significantly robust with respect to background and acceptance effects. In particular, efficiencies common to S_i^+ and S_i^- simplify in the asymmetry calculation, since the latter is invariant under transformations of the type

$$S_i^+ \to \epsilon S_i^+, \qquad S_i^- \to \epsilon S_i^-.$$

6.1.2 Sensitivity to Standard Model and New Physics

In several extensions of the SM, NP effects can modify the Wilson coefficients involved in the $b \to s\ell^+\ell^-$ transition and hence induce a change in the angular observables. In order to examine the expected effect in the angular observables of interest, predictions based on the SM and on several NP scenarios particularly favoured by the current experimental measurements are computed and gathered together in Fig. 6.2. The publicly available `EOS` software [121] and the `BSZ2015` form factors [122] are used in the computation, which is based on QCD factorisation. Three NP scenarios are taken into account: $\mathcal{C}_9^\mu = -1$, $\mathcal{C}_9^\mu = -1.4$, and $\mathcal{C}_9^\mu = -\mathcal{C}_{10}^\mu = -0.64$. The predictions for the $\Delta A_{\rm FB}$ and ΔS_i (with $i = 3, 4, 5, 7, 8, 9$) observables are computed for two different q^2 regions: from $0.1\,\text{GeV}^2/c^4$ to $1.1\,\text{GeV}^2/c^4$ (left) and from $1.1\,\text{GeV}^2/c^4$ to $7.0\,\text{GeV}^2/c^4$ (right). These regions correspond to the q^2 intervals where the measurement described in this thesis is performed and are referred to as *low*

q^2 *bin* and respectively *central* q^2 *bin* in what follows. According to the predictions, no significant discrepancy between SM and NP is expected in the ΔS_3, ΔS_7, ΔS_8, and ΔS_9 observables in any of the two q^2 regions, while the presence of NP can cause a substantial change in the values of the ΔS_4, ΔS_5, and $\Delta A_{\rm FB}$ observables in at least one of the NP scenarios that have been investigated.

6.2 Differences between electrons and muons

The $A_{\rm FB}$ and S_i observables, with $i = 4, 5, 7, 8$, are measured for both the $B^0 \to K^{*0} e^+ e^-$ and $B^0 \to K^{*0} \mu^+ \mu^-$ decays by following a similar procedure for electrons and muons. However, the two particles differ greatly in the way they are detected and reconstructed in the LHCb detector. The reconstruction of both particles relies on the information of the tracking system and the RICH detectors. In addition, muons leave a clear signature in the muon detectors and are reconstructed with high resolution. On the contrary, electrons are reconstructed using the calorimeters, which have a lower resolution than the muon detectors, and must be corrected to account for bremsstrahlung emission, which would otherwise worsen the mass resolution of the B^0 candidates.

If the bremsstrahlung photons are emitted downstream of the magnet, their energy ends up in the same calorimeter cell as the electron and the energy of the electron is reconstructed correctly, provided that no saturation of the calorimeter cell occurs. Conversely, if the photon emission happens upstream of the magnet, the photon and the electron can be detected by two different cells of the calorimeter and the reconstructed energy of the electron does not correspond to the total energy that the particle had initially.

When reconstructing the B^0 candidates, a bremsstrahlung recovery algorithm is used to add the energy of the emitted photons to the electron and positron candidates. The algorithm selects any photon candidate with transverse energy larger than 75 MeV whose position on the ECAL is compatible with the extrapolation of the lepton track from the first tracking station and adds its energy to the energy of the lepton candidate. If the same photon candidate can be associated to both the electron and positron candidates, its energy is added randomly to one of the two. The corrected lepton tracks are then extrapolated backward in order to determine the dilepton vertex and a fit is performed to reconstruct the corrected dilepton momentum. This allows to obtain a better estimate of the momenta and energies of the reconstructed particles, with a corresponding improvement in the mass resolution of the B^0 candidates. The improvement, however, is not sufficient to achieve the same resolution as with decays to muons. This is due to the limitations of the bremsstrahlung recovery algorithm, which arise from the acceptance and resolution of the calorimeters, from the requirement on the minimum transverse energy of the photon candidates that can be associated to the electron and positron candidates, and from the wrong interpretation of energy deposits in the calorimeters as bremsstrahlung photons. As a result, the reconstructed B^0 candidates have a radiative tail at low invariant mass, which is due to the partial recovery of bremsstrahlung photons, and a smaller tail at high invariant mass, which is due to the overcorrection caused by the bremsstrahlung recovery algorithm. These two aspects motivate the invariant mass model used in the measurement, as discussed in Section 7.1.

The worse mass resolution of the B^0 candidates is not the only aspect that makes decays to electrons particularly challenging to investigate. In fact, decays to electrons are also detected with a lower efficiency with respect to the analogous decays to muons. This originates from the lower efficiency of the L0 electron trigger, as described in Section 2.2.4. For this reason, B^0 candidates

triggered by the hadrons (kaons and pions) in the final state or by other activity in the same pp collision are also taken into account in the measurement, as discussed in Section 6.5.

6.3 Samples

This measurement is based on pp collision data, corresponding to an integrated luminosity of 3.0 fb^{-1}, recorded by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. For both $B^0 \to K^{*0} e^+ e^-$ and $B^0 \to K^{*0} \mu^+ \mu^-$ decays, the K^{*0} is reconstructed in the $K^+\pi^-$ final state. This choice is driven by the fact that such a final state is abundant and offers a clear signature due to the presence of two charged particles.

Several simulation samples are used to study the signal and background contributions in the invariant mass, to optimise the selection of the signal candidates, to parametrise the efficiency as a function of q^2 and the decay angles θ_ℓ, θ_K, and ϕ, to validate the counting method approach, and to perform cross-checks of the results obtained in data. In addition to the simulation samples of $B^0 \to K^{*0} e^+ e^-$ and $B^0 \to K^{*0} \mu^+ \mu^-$ to model the rare modes and $B^0 \to K^{*0} J/\psi$ ($\to e^+ e^-$) and $B^0 \to K^{*0} J/\psi$ ($\to \mu^+ \mu^-$) to model the resonant modes, the following decays are simulated:

- $B^+ \to K_1^+ e^+ e^-$;
- $B^+ \to K_2^*(1430)^+ e^+ e^-$;
- $\Lambda_b^0 \to p K^- \ell^+ \ell^-$, with $\ell = e, \mu$;
- $\Lambda_b^0 \to p K^- J/\psi$, with $J/\psi \to \ell^+ \ell^-$ and $\ell = e, \mu$;
- $\Lambda_b^0 \to p \pi^- J/\psi$, with $J/\psi \to \ell^+ \ell^-$ and $\ell = e, \mu$.

The first two simulation samples are used to estimate the background contribution due to *partially reconstructed decays* (see Section 6.7.6). This background contribution must be taken into account in the mass fit of the $B^0 \to K^{*0} e^+ e^-$ candidates, while it is negligible for the analogous decay to muons. The last three simulation samples are used to evaluate the contamination from decays of the Λ_b^0 baryon to final states that can be reconstructed as the rare or control modes (see Section 6.7.2).

The $B^0 \to K^{*0} e^+ e^-$ simulation sample is generated according to the form factors reported in Ref. [123], which are obtained using the technique of the Light Cone Sum Rule (LCSR). The $B^+ \to K_1^+ e^+ e^-$ simulation sample consists of two different modes. The first is a simulation sample with a flat ($K\pi\pi$) invariant mass. The second is a combination of resonant (60%) and non-resonant (40%) decays of the $K_1^+(1720)$ meson, where the resonant component is a mixture of $K^{*0}\pi^+$ (20% of the total), $K^+\rho^0$ (40% of the total), and $K^+\,\omega$ ($< 1\%$ of the total). The $B^+ \to K_2^*(1430)^+ e^+ e^-$ simulation sample is similar to the second $B^+ \to K_1^+ e^+ e^-$ mode but consists of $K_2^*(1430)^+$ instead of $K_1^+(1720)$ decays.

6.4 Dilepton invariant mass

The signal candidates are divided in three exclusive q^2 regions corresponding to the low and central q^2 bins of the rare mode $B^0 \to K^{*0}\, \ell^+ \ell^-$ and a single q^2 bin of the resonant mode $B^0 \to K^{*0} J/\psi$ ($\to \ell^+ \ell^-$). The signal candidates of the rare mode are required to have a q^2 between 0.1 GeV$^2/c^4$ and 1.1 GeV$^2/c^4$ in the low q^2 bin and between 1.1 GeV$^2/c^4$ and 7.0 GeV$^2/c^4$ in

the central q^2 bin. The signal candidates of the resonant mode are instead required to have a q^2 between $7.0\,\text{GeV}^2/c^4$ and $11.0\,\text{GeV}^2/c^4$.

In addition to the requirement on q^2, an equivalent selection is applied to the so-called constrained q^2, hereafter referred to as q^2_c. The latter is computed by constraining the signal candidates to originate from the PV and to have an invariant mass corresponding to the nominal mass of the B^0 meson, as reported in Ref. [15]. This requirement is expected to be very efficient on signal, since the normal and constrained q^2 should have similar values. On the other hand, a good rejection power is expected on background contributions for which the assumption of invariant mass equal to the nominal mass of the B^0 meson does not hold.

The lower bound at $0.1\,\text{GeV}^2/c^4$ allows to easily compare the results of decays to electrons and muons, since both leptons can be considered massless above this threshold. A difference between electrons and muons would hence be a clear hint of LFU violation. The bound between low and central q^2 bins, which is set to $1.1\,\text{GeV}^2/c^4$, is such that the background contribution due to $\phi \to \ell^+\ell^-$ decays is expected to populate the low q^2 bin only. This choice is also consistent with previous measurements at LHCb [34]. The bound between rare and resonant mode, set to $7.0\,\text{GeV}^2/c^4$, is due to the fact that the contamination from the radiative tail of the J/ψ resonance becomes too large above this value. Previous measurements at LHCb considered a q^2 bound of $6.0\,\text{GeV}^2/c^4$ between rare and resonant mode [33, 34]. However, the introduction of the additional requirement on the constrained q^2 allows to move this threshold up to $7.0\,\text{GeV}^2/c^4$, with a significant increase in the statistics available in the rare mode. For the muon mode, the results are given also for the central q^2 bin from $1.1\,\text{GeV}^2/c^4$ to $6.0\,\text{GeV}^2/c^4$, in order to compare with previous measurements at LHCb [35]. For the electron mode, only the results in the nominal central q^2 bin from $1.1\,\text{GeV}^2/c^4$ to $7.0\,\text{GeV}^2/c^4$ are given.

6.5 Trigger requirements

In the offline selection, trigger signals are associated to reconstructed particles. Selection requirements can therefore be made on the trigger selection itself and on whether the decision was due to the signal candidate, other particles produced in the pp collision, or a combination of both. In particular, a signal candidate is classified as:

- *trigger on signal* (TOS), if the trigger signals associated to the signal candidate are sufficient to trigger the event, that is, the event would be triggered even if all trigger signals associated to the rest of the event would be ignored;

- *trigger independent of signal* (TIS), if the trigger signals not associated to the signal candidate are sufficient to trigger the event, that is, the event would still be triggered if the trigger signals associated to the signal candidate are ignored.

According to this definition, a signal candidate can be classified as TIS and TOS at the same time.

In the measurement described here, the signal candidates are required to satisfy the trigger requirements shown in Table 6.1, depending on the presence of muons or electrons in the final state. The conditions for the different trigger levels (L0, HLT1, and HLT2) are combined in `AND`, while those within each trigger level are combined in `OR`. For the electron mode the inclusion of signal candidates triggered by other activity in the same pp collision, corresponding to the `TIS` requirements,

Table 6.1: Trigger requirements for signal candidates with electrons (left) or muons (right) in the final state.

Trigger level	Electron candidates	Muon candidates

L0	`L0Electron TOS`
	`L0Hadron TOS`
	`L0Global TIS`
		`L0Muon TOS`

HLT1	`Hlt1TrackAllL0 TOS`	`Hlt1TrackAllL0 TOS` `Hlt1TrackMuon TOS`

HLT2	`Hlt2Topo2[3,4]BodyBBDT TOS`
	`Hlt2TopoE2[3,4]BodyBBDT TOS`
		`Hlt2Topo2[3,4]BodyBBDT TOS`
		`Hlt2TopoMu2[3,4]BodyBBDT TOS`
		`Hlt2DiMuonDetached TOS`

is needed to increase the statistics available. The requirements of the `L0Electron`, `L0Hadron`, and `L0Muon` trigger lines are described in Section 2.2.4. The `L0Global` trigger line corresponds to the global decision of the L0 trigger. The `Hlt1TrackAllL0` and `Hlt1TrackMuon` trigger lines require at least one good-quality track with high momentum, transverse momentum, and impact parameter χ^2. If the track is associated to hits in the muon detectors (`Hlt1TrackMuon`), the requirements on momentum, transverse momentum, and impact parameter χ^2 are relaxed [124]. The topological lines `Hlt2Topo2[3,4]BodyBBDT`, `Hlt2TopoE2[3,4]BodyBBDT`, and `Hlt2TopoMu2[3,4]BodyBBDT` are described in Section 2.2.4 and differ in the number and type of charged particles reconstructed in the final state. The `Hlt2DiMuonDetached` trigger line selects dimuon resonances with masses below the J/ψ mass. The two tracks associated to the muons are required to have large impact parameter χ^2, while the dimuon candidate is required to have a mass larger than $1\,\text{GeV}/c^2$ and a transverse momentum larger than $1.5\,\text{GeV}/c$ [124].

Since the distribution of the signal candidates in the electron mode depends on the nature of the particle that has triggered the event (lepton or hadron), three independent trigger categories are taken into account. These are defined as follows:

- `L0 Electron` (`L0E`) contains signal candidates that are `TOS` with respect to the `L0 Electron` trigger line;

- `L0 Hadron` (`L0H`) contains signal candidates that are `TOS` with respect to the `L0 Hadron` trigger line and do not belong to the above category;

- `L0 TIS` (`L0TIS`) contains signal candidates that are `TIS` with respect to the `L0 Global` trigger line (that is, the corresponding event has been triggered by other particles) and do not belong to one of the two categories above.

The measurement of the S_i observables of the resonant mode $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ is performed separately in each of these categories, analogously to what described in Ref. [34]. A different approach is adopted for the measurement of the S_i observables of the rare mode $B^0 \to K^{*0} e^+e^-$. Since the statistics available does not allow to perform a mass fit separately for each trigger category, only one fit is performed.

6.6 Corrections to simulation

The simulation is corrected to account for some known discrepancies between simulation and data. These discrepancies originate from a mismodelling of the kinematics of the event, as well as from a mismodelling of the detector response in terms of particle identification and trigger efficiency.

The corrections consist of three steps: the first corrects the response of the particle identification algorithm, the second corrects the modelling of the kinematics of the event, and the third corrects the response of the trigger algorithm, *i.e.* the trigger efficiency. Each step is described in detail in the next subsections. However, an outline of the corrections is anticipated here in order to provide the reader with a general understanding of the procedure. In the first step, a data-driven approach is used to resample the PID distributions of the simulated candidates, so that the resampled variables can be used in the selection. In the second step, a multidimensional reweighting based on a multivariate technique allows to assign event-by-event weights to the simulated candidates. These weights, w_{BDT}, are normalised to the total number of simulated candidates. In the third step, the w_{BDT} weights are taken into account to assign a correction factor, w_{trig}, to each simulated candidate. As a result of this procedure, the weighted distributions of the simulation reproduce the behaviour observed in data. This is needed since the simulation is used as a proxy of the signal and background contributions in several parts of the analysis, for example in the multivariate analysis described in Section 6.7.5.

6.6.1 Particle identification

The particle identification algorithm used in LHCb is based on a multivariate classifier that takes the response of the PID detectors into account. The latter depends on detector occupancy and operating conditions of the detectors, which can often change within a given data taking period. Both aspects are extremely difficult to model and cause a disagreement of the response of the particle identification algorithm between simulation and data.

A common procedure that allows to mitigate this discrepancy is to follow a data-driven approach to resample the PID distributions of the simulated candidates. In order to do so, high-purity calibration samples are used to determine the distribution of a given PID variable x in both simulation and data as a function of a set of key variables. For example, a common choice in LHCb is to parametrise the distribution of the PID variable x as a function of the number of tracks in the pp collision, n_{tracks}, and two kinematic variables, *e.g.* the transverse momentum, p_{T}, and the pseudorapidity, η, of the particle to which the PID variable x is referred. The probability density function (PDF) observed in data can hence be written as $p_{\text{exp}}(x|p_{\text{T}}, \eta, n_{\text{tracks}})$. The corresponding cumulative distribution function (CDF) is

$$P_{\text{exp}}(x|p_{\text{T}}, \eta, n_{\text{tracks}}) = \int_{-\infty}^{x} p_{\text{exp}}(y|p_{\text{T}}, \eta, n_{\text{tracks}})\, dy.$$

This formula can be used to associate a new PID variable $x_{\text{corr}} = P^{-1}_{\text{exp}}(\xi|p_{\text{T}}, \eta, n_{\text{tracks}})$ to the simulated candidates. If the sampling variable ξ is distributed uniformly between zero and unity, the distribution of x_{corr} is $p_{\text{exp}}(x_{\text{corr}}|p_{\text{T}}, \eta, n_{\text{tracks}})$ and hence reproduces what observed in data. A slightly different approach is followed in this measurement: the sampling variable ξ is not chosen

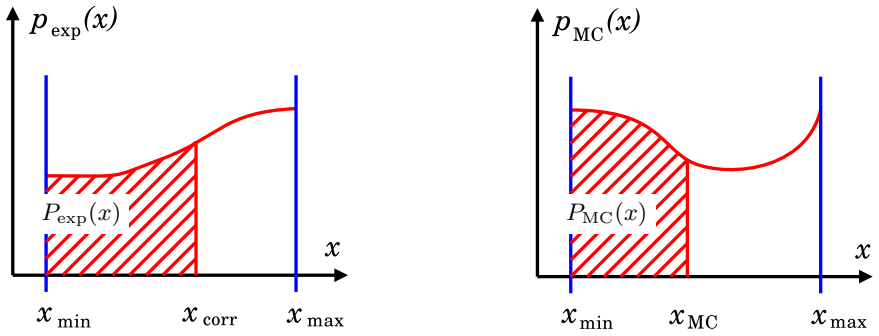


Figure 6.3: Schematic view of the PDF obtained from data (left) and simulation (right). The transformation of variables $x_{\text{MC}} \to x_{\text{corr}}$ is obtained by requiring $P_{\text{exp}}(x_{\text{corr}}) = P_{\text{MC}}(x_{\text{MC}})$.

randomly but is obtained from the original PID variable x_{MC} as follows:

$$\xi = P_{\text{MC}}(x_{\text{MC}}|p_{\text{T}}, \eta, n_{\text{tracks}}) = \int_{-\infty}^{x_{\text{MC}}} p_{\text{MC}}(y|p_{\text{T}}, \eta, n_{\text{tracks}})\, dy,$$

where $p_{\text{MC}}(x|p_{\text{T}}, \eta, n_{\text{tracks}})$ is the PDF observed in simulation. Therefore, the transformation of variables $x_{\text{MC}} \to x_{\text{corr}}$ is obtained by requiring

$$x_{\text{corr}} = P_{\text{exp}}^{-1}\left(P_{\text{MC}}(x_{\text{MC}}|\, p_{\text{T}}, \eta, n_{\text{tracks}})\,|\, p_{\text{T}}, \eta, n_{\text{tracks}}\right).$$

This is shown schematically in Fig. 6.3. This method has the advantage of preserving the correlation between the original PID variable x_{MC} and the new PID variable x_{corr}, at least if $p_{\text{exp}}(x)$ and $p_{\text{MC}}(x)$ do not differ too much from each other. As a consequence, the correlation between the PID variable x and the other properties of the event not involved in the resampling procedure is also partially preserved.

In this measurement, the distributions of the DLL and ProbNN variables associated to the particles in the final state (see Section 6.7.1 for a more detailed description of how these variables are defined) are corrected using three classes of high-purity control samples: $D^{*+} \to D^0\pi^+$, $\Lambda^0 \to p\pi^-$, and $\Lambda_c^+ \to pK^+\pi^-$ decays are used to correct the PID distributions of hadrons, $J/\psi \to \mu^+\mu^-$ decays are used to correct muons, and $B^+ \to K^+J/\psi\,(\to e^+e^-)$ are used to correct electrons. The functional dependence of $p_{\text{exp}}(x|p_{\text{T}}, \eta, n_{\text{tracks}})$ is obtained separately for 2011 and 2012 and for the two magnet polarities by following the `Meerkat` method described in Ref. [125]. This method has been proposed recently in LHCb and is based on a kernel density estimation (KDE) technique. Starting from a calibration sample of data points y_i, where y is a single variable or a vector of variables, the method allows to estimate the PDF $p(y)$ without the need to know the physical model behind the process under study.

The following transformations of variables are applied to improve the uniformity of the distributions entering the KDE calculation:

$$p_{\text{T}}{}' = \ln p_{\text{T}}, \quad n'_{\text{tracks}} = \ln n_{\text{tracks}},$$

for the variables used in the parametrisation of the PID distributions and

$$\text{ProbNN}_\text{x}{}' = 1 - (1 - \text{ProbNN}_\text{x})^\gamma, \quad \text{ProbNN}_\text{x}{}' = \text{ProbNN}_\text{x}{}^\gamma, \tag{6.2}$$

with $\gamma = 1/2$, for the ProbNN variables. The first transformation of Eq. (6.2) is used for the PID variables referring to the correct PID hypothesis (*e.g.* the PID variable ProbNN_e of the electron candidate), while the second is used for those referring to the wrong PID hypothesis (*e.g.* the PID variable ProbNN_e of the kaon candidate). The different transformations of variables are due to the fact that the corresponding distributions peak at unity and zero, respectively. No transformation of variables is instead required for the DLL variables.

An example of the original and resampled PID distributions of some of the key variables used in the analysis is shown in Fig. 6.4 for the $B^0 \to K^{*0} J/\psi$ $(\to e^+e^-)$ simulated candidates. The corresponding PID distributions observed in data are overlaid for comparison. A good agreement is observed over the entire range of all the resampled PID variables. Any further residual differences between simulation and data are accommodated as systematic uncertainties as discussed in Section 7.4.

6.6.2 Kinematics of the event

The mismodelling of the kinematics of the event is corrected by performing a multidimensional reweighting of the variables that show the largest deviation between simulation and data. The multidimensional reweighting is based on the method described in Ref. [126] and is referred to as Boosted Decision Tree (BDT) reweighting in the following. This method has several advantages compared to the standard histogram reweighting. In the latter, the distributions of the reweighting variables in simulation (*original sample*) and data (*target sample*) are split in bins and the bin-by-bin ratio

$$w_i = w_i^{\text{target}} / w_i^{\text{original}}$$

is used as weight of the simulated candidates. In the formula, w_i^{target} and w_i^{original} represent the total weight of the events in the i^{th} bin of the target and the original distribution, respectively. Despite its simplicity, the histogram reweighting has some limitations. First of all, it cannot be used to reweight many variables at the same time, since the statistics in each multidimensional bin would be too low. In addition, even if the *reweighted sample* and the target sample are in perfect agreement with each other, there is no guarantee that the procedure will work well on a different simulation sample. In the BDT reweighting the first limitation is overcome by the fact that the space of the reweighting variables is split in regions by using decision trees. The optimal regions are identified by maximising the symmetrised binned χ^2 defined as

$$\chi^2 = \sum_{\text{reg}} \frac{\left(w^{\text{reg}}_{\text{original}} - w^{\text{reg}}_{\text{target}}\right)^2}{w^{\text{reg}}_{\text{original}} + w^{\text{reg}}_{\text{target}}},$$

where $w^{\text{reg}}_{\text{original}}$ and $w^{\text{reg}}_{\text{target}}$ are the total weight of the events in the selected region of the target and the original distribution, similarly to what discussed above.

The performance of the reweighting procedure is evaluated by training a multivariate classifier to distinguish between simulation and data before and after the BDT reweighting is applied. If the

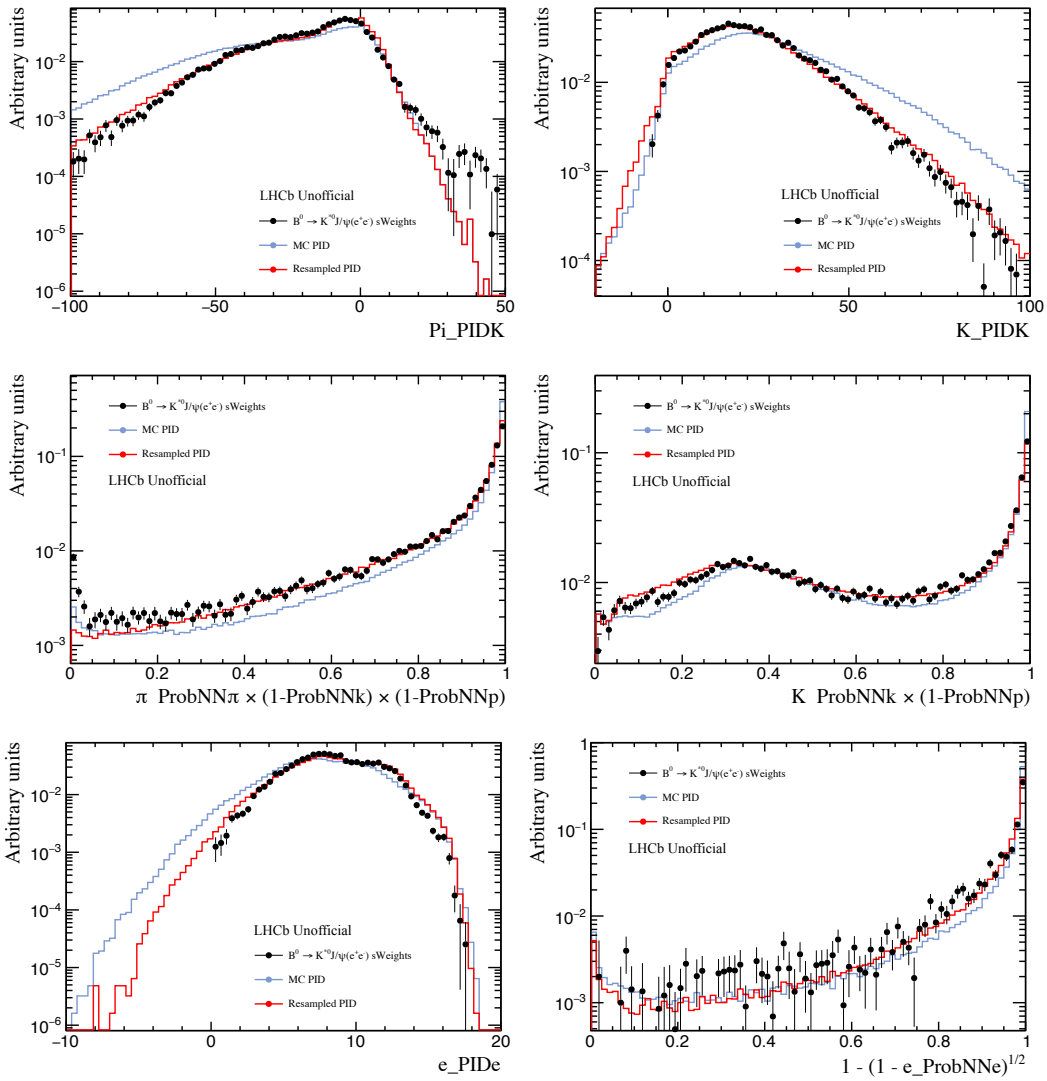


Figure 6.4: Distributions of a representative set of PID variables used in the analysis obtained from $B^0 \to K^{*0}J/\psi$ ($\to e^+e^-$) simulated candidates before (blue solid line) and after (red solid line) the PID correction is applied. The corresponding PID distributions observed in *sWeighted* data are overlaid for comparison (black points with error bars).

reweighter improves the agreement between simulation and data, in fact, the discrimination between the two becomes harder when the correction is applied. The performance of this algorithm can be checked by comparing two different trainings of the same multivariate classifier, one based on the original sample and target sample and another based on the reweighted sample and target sample.

The BDT reweighting procedure adopted in this measurement consists of two subsequent steps: a first BDT reweighting obtained from the muon sample and a second BDT reweighting obtained from either the muon sample or the electron sample, after the latter are corrected with the weights obtained in the first step. This allows to account for potential correlations among the features (that is, the variables) used in the two steps. Eventually, a weight is assigned to each simulated candidate, corresponding to the combined correction obtained from the two steps. The first BDT reweighting allows to correct the mismodelling of the kinematics of the event independently of the

trigger selection, while the second allows to correct the residual discrepancies induced by the trigger selection applied to the samples. For the first step, an inclusive trigger category, where electrons and muons are expected to be selected in a similar matter, is considered. This is achieved by requiring that the B^0 candidates are selected due to the activity of the rest of the event, which corresponds to the following trigger requirement:

$$\texttt{L0 Global} \text{ TIS} \cap \texttt{Hlt1 Phys} \text{ TIS} \cap \texttt{Hlt2 Phys} \text{ TIS}.$$

This guarantees that the corresponding sample is not biased by the trigger selection. For the second step, the trigger selection described in Section 6.5 is applied.

The $B^0 \to K^{*0} J/\psi$ $(\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi$ $(\to e^+e^-)$ decay modes are used as input for the multidimensional reweighting, since a clear signal is visible in data after the preselection described in 6.7.1 and the requirements to decrease the contamination from exclusive backgrounds described in 6.7.2 are applied. In addition, the PV- and J/ψ-constrained invariant mass of the B^0 candidates (that is, the invariant mass computed by constraining the B^0 candidates to originate from the PV and the J/ψ candidates to have an invariant mass corresponding to the nominal mass of the J/ψ, as reported in Ref. [15]) is required to be larger than $4800\,\text{MeV}/c^2$. Only B^0 candidates with successful kinematic fit when the two constraints are taken into account are retained. For the muon mode, the invariant mass of the J/ψ candidates is required to be within $100\,\text{MeV}/c^2$ from the nominal value, while a looser selection of q^2 between $6.0\,\text{GeV}^2/c^4$ and $11.0\,\text{GeV}^2/c^4$ is applied to the electron mode.

The *sPlot* technique [127] is used to statistically subtract the background and determine the distribution of the signal candidates in data. This is achieved by fitting the PV- and J/ψ-constrained invariant mass distribution of the B^0 candidates. The fit is performed separately for 2011 and 2012. The model used in the fit to the data samples is partially obtained from simulation. The $B^0 \to K^{*0} J/\psi$ $(\to \ell^+\ell^-)$ simulated candidates are fitted with a Double Crystal Ball (DCB), which consists of two Crystal Balls (CBs) with shared mean:

$$\text{DCB} = f_{\text{CB}}\text{CB}_1 + (1 - f_{\text{CB}})\text{CB}_2.$$

Each CB consists of a Gaussian distribution with power-law tails and is parametrised as follows:

$$\text{CB}\,(x;\alpha,n,\mu,\sigma) = N \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}} & \text{if } \frac{x-\mu}{\sigma} > -\alpha \\ A\left(B - \frac{x-\mu}{\sigma}\right)^{-n} & \text{if } \frac{x-\mu}{\sigma} \leq -\alpha \end{cases}$$

where N, A, and B are normalisation constants defined as:

$$N = \frac{1}{\sigma(N_1+N_2)} \quad \text{with } N_1 = \frac{n}{(n-1)|\alpha|}e^{-\frac{|\alpha|^2}{2}} \text{ and } N_2 = \sqrt{\frac{\pi}{2}}\left(1 + \text{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right)$$

and

$$\begin{cases} A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}} \\ B = \frac{n}{|\alpha|} - |\alpha| \end{cases}$$

The error function is approximated numerically as described in Ref. [128]. The values of f_{CB}, α_1, α_2, n_1, and n_2 obtained from the fit to the simulation samples are used as constants in the fit to the data samples. The sum of two DCBs and an exponential distribution is used for the unbinned

extended maximum likelihood fit of the $B^0 \to K^{*0} J/\psi$ $(\to \ell^+\ell^-)$ data candidates:

$$\mathrm{PDF}_{\mathrm{data}} = \mathcal{N}_{B^0}\ \mathrm{DCB}_{B^0} + \mathcal{N}_{B^0_s}\ \mathrm{DCB}_{B^0_s} + \mathcal{N}_{\mathrm{comb\ bkg}}\ \mathrm{EXP}_{\mathrm{comb\ bkg}},$$

where the two DCBs model the $B^0 \to K^{*0} J/\psi$ $(\to \ell^+\ell^-)$ and $B^0_s \to K^{*0} J/\psi$ $(\to \ell^+\ell^-)$ contributions and the exponential distribution parametrises the combinatorial background. In order to decrease the degrees of freedom of the total PDF and improve the stability of the fit, the two DCBs have all parameters in common except for the mean. For the B^0_s DCB, the latter is fixed to the mean of the B^0 DCB plus the difference between the nominal masses of the B^0_s and B^0 mesons, as reported in Ref. [15]. This allows to account for momentum scale effects of the detector, which might shift the measured value of the masses, with a similar shift expected for the two particles. The yields of the three components of the model, $\mathcal{N}_{B^0}$, $\mathcal{N}_{B^0_s}$, and $\mathcal{N}_{\mathrm{comb\ bkg}}$, are left free to vary in the fit, together with the mean μ, the widths σ_1 and σ_2, and the slope of the exponential distribution. The fit results are shown in Figs. 6.5 and 6.6 for the muon and the electron mode, respectively.

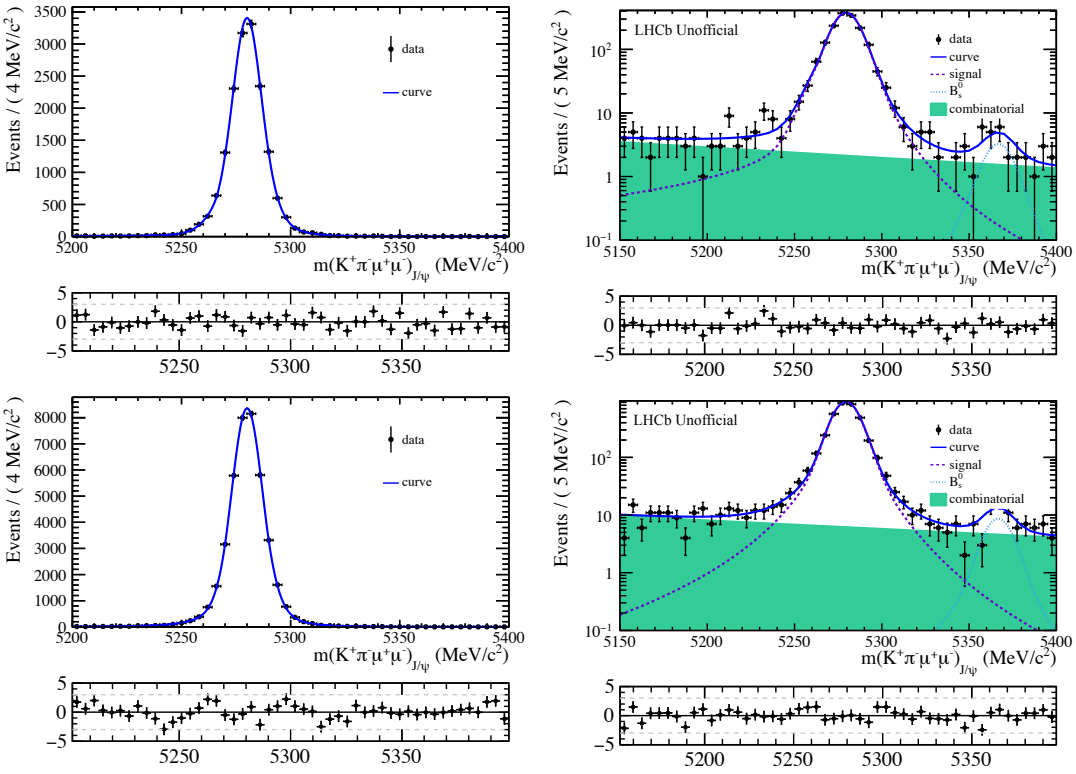


Figure 6.5: Distribution of the $B^0 \to K^{*0} J/\psi$ $(\to \mu^+\mu^-)$ candidates (black points with error bars) and corresponding fit (solid blue line) for 2011 (top) and 2012 (bottom) simulation (left) and data (right). The fit model uses a DCB for simulation and two DCBs and an exponential distribution for data. The residuals are also shown.

The differences between the distributions of the *sWeighted* data candidates and the simulated candidates are used to develop the BDT reweighting procedure. In the first step, the simulation distributions of the number of tracks, number of SPD hits, B^0 p_{T}, and B^0 decay vertex χ^2/ndf are reweighted to reproduce the distributions observed in data. In the second step, the BDT reweighting is trained on the residual discrepancies in the minimum and maximum transverse momentum and

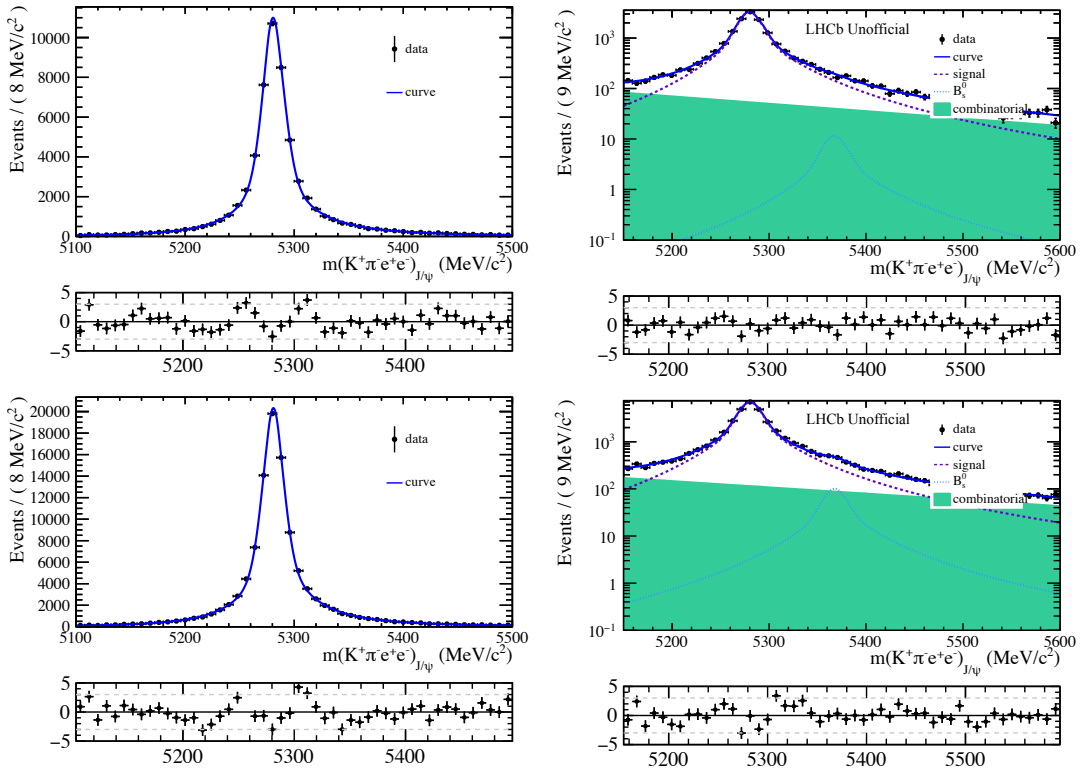


Figure 6.6: Distribution of the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ candidates (black points with error bars) and corresponding fit (solid blue line) for 2011 (top) and 2012 (bottom) simulation (left) and data (right). The fit model uses a DCB for simulation and two DCBs and an exponential distribution for data. The residuals are also shown.

pseudorapidity of the leptons and hadrons in the final state. This is done separately for muons and electrons. Since the BDT reweighting is performed separately for 2011 and 2012, the differences in the p_{T} distributions caused by the different collision energy in the two years are correctly taken into account.

The BDT reweighting is based on the `raredecay` package [129] and uses two `GBReweighter` reweighters, which share the same configuration apart from the number of estimators. In order to make a better use of the statistics of the samples used for the training of the BDT reweighters, a k-fold cross-validation approach is adopted, where 9/10 of the data are used to determine the correction, which is then applied to the remaining 1/10. In order to increase the stability of the results, the BDT reweighters are applied to each signal candidate 20 times and the average weight is then used. The performance of the BDT reweighting is compared to that achievable using a standard histogram reweighting with 30 bins in the first step and 10 bins in the second and is found to be superior to it. All variables used in the multivariate analysis described in Section 6.7.5, as well as the decay angles θ_ℓ, θ_K, and ϕ and other variables relevant for the analysis are taken into account to score the performance of the BDT reweighting. The score is computed using an `XGBoost` classifier.

In order to simpify the procedure, an alternative strategy in which the number of tracks is not included in the subset of variables used in the first BDT reweighting is also investigated. It is observed that the exclusion of the number of tracks causes an overcorrection of the remaining variables. In particular, the distribution of the number of SPD hits after the correction is applied

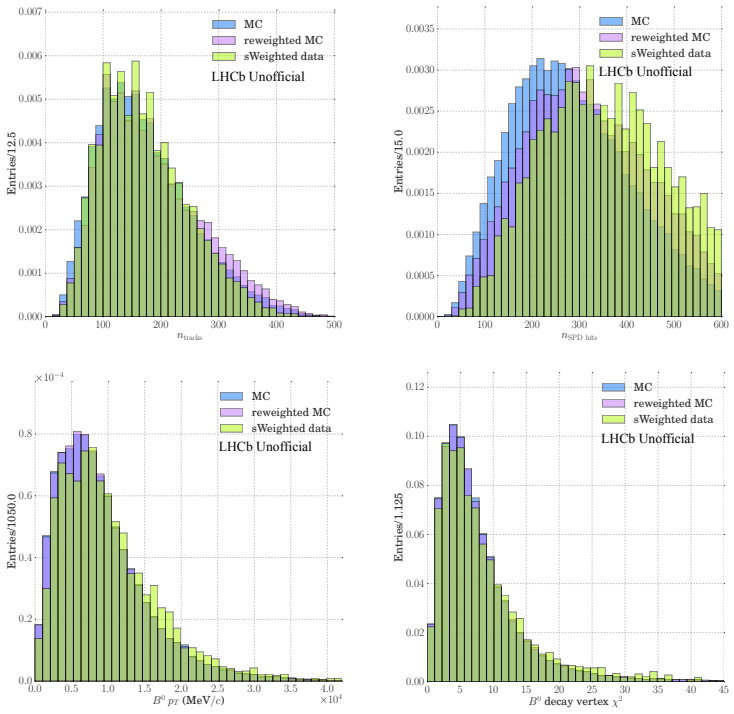


Figure 6.7: Distribution of the number of tracks (top left), number of SPD hits (top right), B^0 p_{T} (bottom left), and B^0 decay vertex χ^2/ndf (bottom right) for $B^0 \to K^{*0} J/\psi$ $(\to \mu^+\mu^-)$ *sWeighted* data (green) and simulation before (blue) and after (violet) the correction is applied. The distributions are obtained from 2012 samples.

shows a better similarity between simulation and data compared to the case in which the number of tracks is used in the training. However, the agreement between the simulation and data distributions of the number of tracks after the correction is applied gets worse than the original. For this reason, it is preferred to adopt the strategy including the number of tracks.

The improvement in the agreement between data and simulation can be seen in Figs. 6.7 and 6.8, which show the comparison between data and simulation distributions before and after the correction is applied. The distributions in Figs. 6.7 and 6.8 are obtained from 2012 samples. Similar results are observed for 2011. The distributions that undergo the major changes due to the BDT reweighting are the number of tracks and the number of SPD hits, while the changes of the other distributions are mostly marginal. The reweighted distributions do not match exactly the target distributions since a stronger reweighting of the selected variables would induce a disagreement in some other variables used in the analysis.

6.6.3 Trigger efficiency

The trigger efficiency, that is, the efficiency of triggering the signal candidates that were reconstructed by the detector, is not well described in the simulation. Therefore, a correction factor is assigned to each simulated candidate, so that the simulation distribution of the trigger efficiency reproduces the trend observed in data. This is performed in two subsequent steps: a first correction factor is used to correct the L0 trigger efficiency and a second correction factor, obtained after applying the first

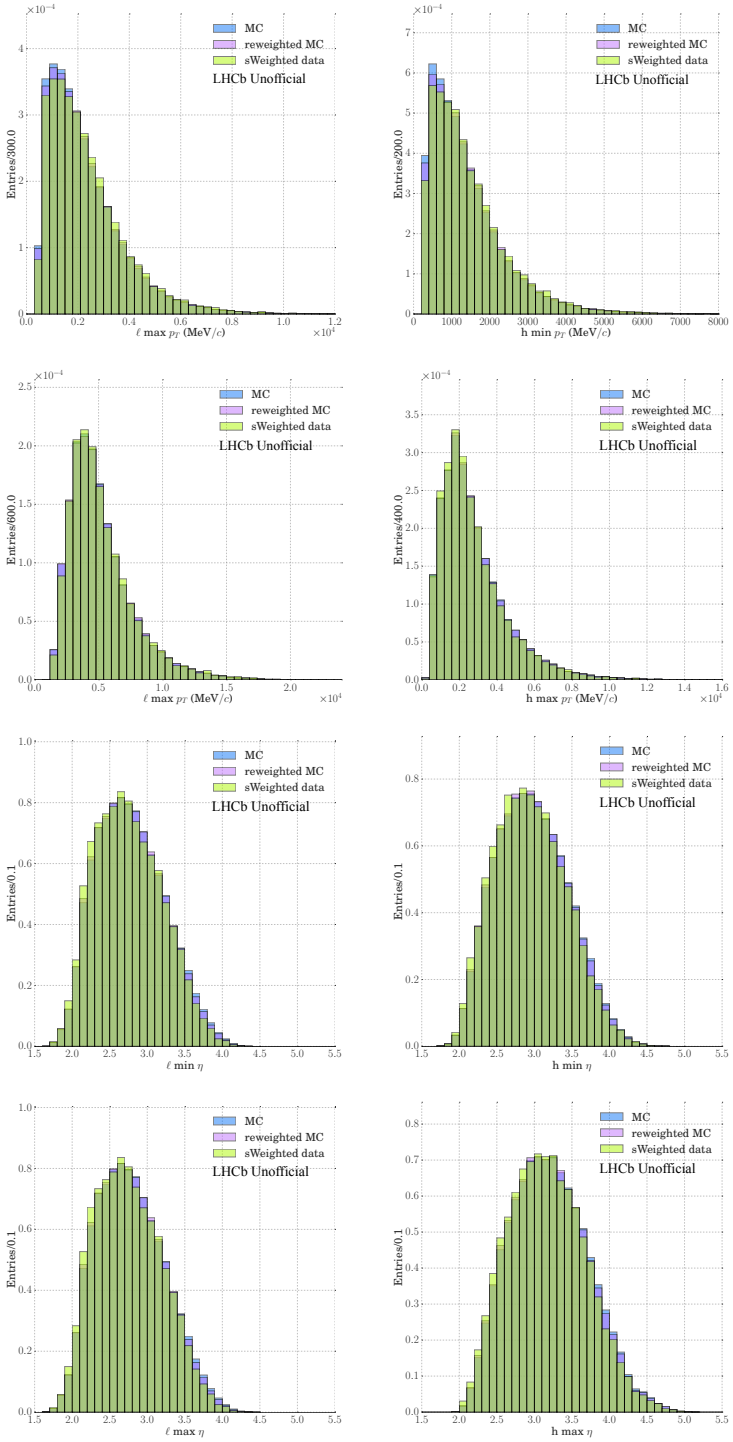


Figure 6.8: Distribution of the minimum transverse momentum (first row), maximum transverse momentum (second row), minimum pseudorapidity (third row), and maximum pseudorapidity (fourth row) of leptons (left) and hadrons (right) for $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) *sWeighted* data (green) and simulation before (blue) and after (violet) the correction is applied. The distributions are obtained from 2012 samples.

correction, is used to correct the HLT trigger efficiency. For each step, the correction factor used to correct the simulated candidates corresponds to the ratio between data and simulation trigger efficiencies:

$$w(\vec{x}) = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}},$$

where the efficiencies are computed as a function of one or two selected variables, generically referred to as $\vec{x}$. For the electron mode, this is done separately for the three trigger categories `L0E`, `L0H`, and `L0TIS` in which the analysis is performed. The weights obtained in Section 6.6.2 are taken into account when determining the trigger efficiency of the simulation samples.

The trigger efficiencies ϵ_{data} and ϵ_{MC} are computed by using a tag and probe method, referred to as TISTOS in the following. This method uses a reference sample that is independent of the trigger selection, the TIS sample, and counts how many signal candidates in this reference sample are compatible with causing a positive trigger decision, that is, are TOS with respect to the relevant trigger. Formally, this corresponds to computing the ratio

$$\epsilon = \frac{\mathcal{N}_{\text{TOS}\wedge\text{TIS}}}{\mathcal{N}_{\text{TIS}}},$$

where $\mathcal{N}_{\text{TIS}}$ is the number of signal candidates in the reference sample and $\mathcal{N}_{\text{TOS}\wedge\text{TIS}}$ is how many of these signal candidates are compatible with triggering.

The choice of the TIS and TOS samples depends on the specific correction under study. Similarly to the correction of the kinematics of the event described in Section 6.6.2, the resonant mode $B^0 \to K^{*0} J/\psi$ $(\to \ell^+\ell^-)$ (both simulation and data) is used as input to evaluate the trigger efficiency correction. Instead of using the *sPlot* technique to obtain background-subtracted data samples, $B^0 \to K^{*0} J/\psi$ $(\to \ell^+\ell^-)$ data candidates are required to have a PV- and J/ψ-constrained B^0 invariant mass within $\pm 60\,\text{MeV}/c^2$ from the nominal value of the B^0 mass. As for the correction of the kinematics of the event described in 6.6.2, the invariant mass of the J/ψ candidates is required to be within $100\,\text{MeV}/c^2$ from the nominal value, in the muon mode, and between $6.0\,\text{GeV}^2/c^4$ and $11.0\,\text{GeV}^2/c^4$, in the electron mode. The correction factors are obtained separately for 2011 and 2012, since the trigger requirements were slightly different in the two years.

The corrections for the signal candidates in the `L0M` and `L0E` trigger categories obtained for 2012 samples are shown as an example in Fig. 6.9. For the `L0M` trigger category, the L0 correction factor w_{L0} is obtained as a function of the maximum transverse momentum of the leptons in the final state, as shown in Fig. 6.9 (top left), while the HLT correction factor w_{HLT} is obtained as a function of the minimum transverse momentum of the particles in the final state, as shown in Fig. 6.9 (bottom left). The different choice of variables is due to the fact that the trigger selection of the muon sample is based on the presence of high-momentum particles, at the L0 level, and on several variables, including the presence of low-momentum particles, at the HLT level. For the `L0E` trigger category, the L0 correction factor is obtained as a function of the maximum transverse energy of the leptons in the final state and the region of the ECAL where the corresponding energy cluster is located (inner, middle, or outer region). The correction corresponding to the outer region is shown in Fig. 6.9 (top right). The HLT correction factor is instead obtained following the same procedure of the `L0M` trigger category. The L0 and HLT corrections for the signal candidates in the `L0H` and `L0TIS` trigger categories are obtained in a similar way, but are not described further here.

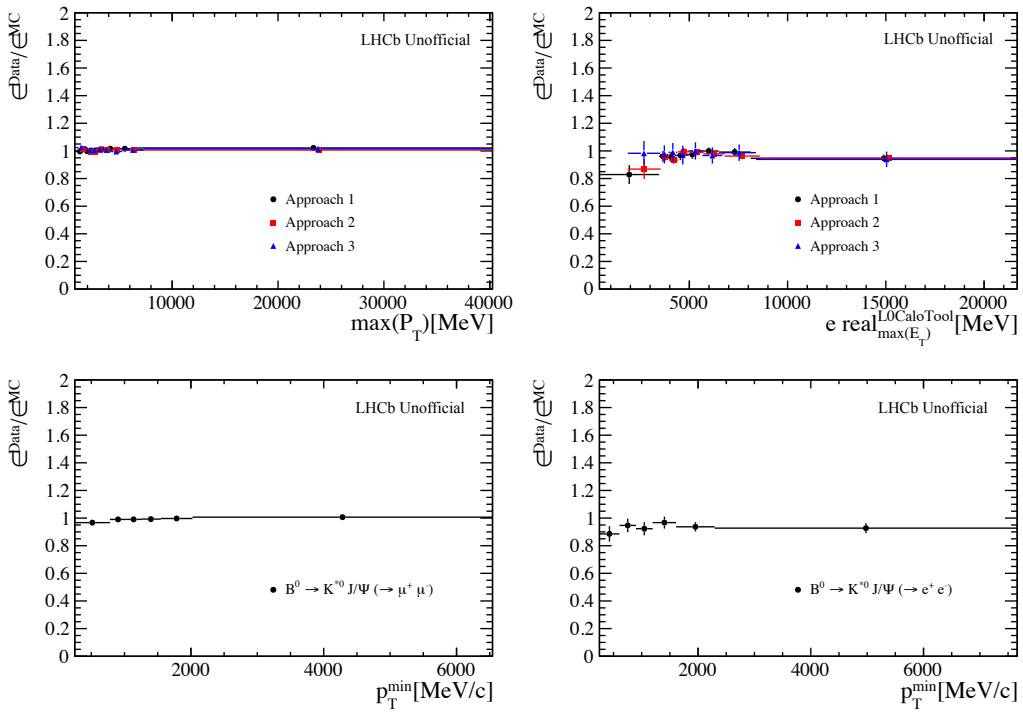


Figure 6.9: Ratio between data and simulation trigger efficiencies used to correct the L0 (top) and HLT (bottom) trigger efficiency of the `L0M` (left) and `L0E` (right) simulation samples. The plots showing the L0 trigger efficiency are obtained from leptons leaving an energy cluster in the outer region of the ECAL. Similar plots are available for the inner and middle regions of the ECAL, as well as for the `L0H` and `L0TIS` trigger categories. The three colors in the plots of the L0 trigger efficiency refer to three alternative approaches that are used to evaluate the trigger efficiency.

6.7 Selection

6.7.1 Preselection

A B^0 candidate is formed from a pair of well-reconstructed oppositely-charged particles identified as either two muons or two electrons and a pair of well-reconstructed oppositely-charged particles, one identified as a kaon and the other identified as a pion.

The complete list of preselection requirements applied to simulation and data samples is shown in Table 6.2. Candidates belonging to pp collisions without any reconstructed PV or with more than 600 hits in the SPD are removed from the samples. Particles in the final state are required to have ghost probability smaller than 0.4, that is, low probability of originating from fake tracks, and track χ^2/ndf smaller than 3, which allows to select good-quality tracks. The kaon, pion, electron, and muon candidates are required to have associated hits in the RICH detectors. Electrons (muons) are also required to have associated clusters in the calorimeters (hits in the muon detectors). The particles in the final state are required to be compatible with their correct PID hypothesis. In order to do so, two types of PID variables are used in the selection. Both types are based on a multivariate classifier which takes the information of the PID detectors into account. The first type, referred to as $\mathrm{DLL}_{a\pi}$, with a identifying a particle type other than π, corresponds to the difference between the

log-likelihood of the particle to be a or to be π:

$$\mathrm{DLL}_{a\pi} = \log \mathcal{L}_a - \log \mathcal{L}_\pi.$$

The second type, referred to as ProbNN_b, corresponds to the response of an Artificial Neural Network (ANN) that uses the information of the tracking and particle identification systems as input. The response of the ANN is normalised between zero and one and does not take the pion PID hypothesis as reference. In the measurement described here, kaon candidates are required to have $\mathrm{DLL}_{K\pi}$ larger than -5, while electron candidates are required to have $\mathrm{DLL}_{e\pi}$ larger than 0. In addition, the following ProbNN requirements are applied:

- for kaon candidates, $\mathrm{ProbNN}_K \cdot (1 - \mathrm{ProbNN}_p) > 0.05$;
- for pion candidates, $\mathrm{ProbNN}_\pi \cdot (1 - \mathrm{ProbNN}_K) \cdot (1 - \mathrm{ProbNN}_p) > 0.1$;
- for electron candidates, $\mathrm{ProbNN}_e > 0.2$;
- for muon candidates, $\mathrm{ProbNN}_\mu > 0.2$.

A minimum p_T of 250 MeV/c is required for kaons and pions, while a threshold value of 500 (800) MeV/c is required for electrons (muons). All particles in the final state are required to have large impact parameter χ^2 with respect to the PV. In addition, quality requirements, like having a good-quality vertex, are applied to the dilepton and $K\pi$ pairs, as well as to the B^0 candidate. The B^0 candidate is required to be compatible with originating from the PV, while the dilepton and $K\pi$ pairs are required to be significantly displaced from it, as expected given the long lifetime of the B^0 from which they originate. The invariant mass of the dilepton pair is required to be smaller than 5500 MeV/c^2, while the invariant masses of the $K\pi$ pair and of the B^0 candidate are required to be respectively within 100 MeV/c^2 and 1000 MeV/c^2 from their nominal values. The requirement on the invariant mass of the $K\pi$ pair is relaxed to 200 MeV/c^2 instead of 100 MeV/c^2 in some parts of the analysis (see for example Section 6.7.5) in order to increase the statistics available. A minimum p_T of 500 MeV/c is required for the $K\pi$ pair.

6.7.2 Exclusive backgrounds

In addition to the requirements described in the previous section, specific requirements are applied to lower the contamination from the exclusive decays $B_s^0 \to \phi\ \ell^+\ell^-$, $B^+ \to K^+\ \ell^+\ell^-$, $\Lambda_b^0 \to pKJ/\psi\ (\to \ell^+\ell^-)$, $\Lambda_b^0 \to pK\ \ell^+\ell^-$, and $B^0 \to D^-\ (\to K^{*0}\ell^-\overline{\nu}_\ell)\ell^+\nu_\ell$. Further requirements are applied to remove another source of background, which originates from hadron-lepton and hadron-hadron swaps, that is, from signal decays with $K \to \ell$, $\pi \to \ell$, $K \to \pi$, or $\pi \to K$ misidentification. In order to keep the selection of electron and muon decays as similar as possible, the same requirements are applied to both. A selection to reduce the contamination from partially reconstructed decays is also discussed.

$B_s^0 \to \phi\ell^+\ell^-$ background

If a kaon from the $\phi \to K^+K^-$ decay is misidentified as pion, the ϕ of a $B_s^0 \to \phi\ell^+\ell^-$ decay can be reconstructed as K^{*0} and end up in the selected K^{*0} invariant mass window. This source of

Table 6.2: Preselection requirements.

Object	Requirement
Event	$N_{\mathrm{PV}} \geq 1$
	$n_{SPD} \leq 600$
K	track $\chi^2/\mathrm{ndf} < 3$
	track ghost probability < 0.4
	`hasRICH`
	$\mathrm{DLL}_{K\pi} > -5$
	$\mathrm{ProbNN}_K \cdot (1 - \mathrm{ProbNN}_p) > 0.05$
	$p_{\mathrm{T}} > 250\,\mathrm{MeV}/c$
	χ^2_{IP} (primary) > 9
π	track $\chi^2/\mathrm{ndf} < 3$
	track ghost probability < 0.4
	`hasRICH`
	$\mathrm{ProbNN}_\pi \cdot (1 - \mathrm{ProbNN}_K) \cdot (1 - \mathrm{ProbNN}_p) > 0.1$
	$p_{\mathrm{T}} > 250\,\mathrm{MeV}/c$
	χ^2_{IP} (primary) > 9
e	track $\chi^2/\mathrm{ndf} < 3$
	track ghost probability < 0.4
	`hasRICH`
	`hasCalo`
	$\mathrm{DLL}_{e\pi} > 0$
	$\mathrm{ProbNN}_e > 0.2$
	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$
	χ^2_{IP} (primary) > 9
μ	track $\chi^2/\mathrm{ndf} < 3$
	track ghost probability < 0.4
	`hasRICH`
	`isMuon`
	$\mathrm{ProbNN}_\mu > 0.2$
	$p_{\mathrm{T}} > 800\,\mathrm{MeV}/c$
	χ^2_{IP} (primary) > 9
$\ell\ell$	$m < 5500\,\mathrm{MeV}/c^2$
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 9$
	PV χ^2 separation > 16
K^{*0}	$\lvert m - m^{\mathrm{PDG}}_{K^{*0}}\rvert < 100\,\mathrm{MeV}/c^2$
	$p_{\mathrm{T}} > 500$ MeV
	χ^2_{IP} (primary) > 9
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 25$
B^0	$\lvert m - m^{\mathrm{PDG}}_{B^0}\rvert < 1000\,\mathrm{MeV}/c^2$
	`DIRA` > 0.9995
	χ^2_{IP} (primary) < 25
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf} < 9$
	PV χ^2 separation > 100

background is reduced by requiring the invariant mass of the $K\pi$ pair, calculated assigning the kaon hypothesis to the pion, to be larger than 1040 MeV/c^2:

$$m_{(K(\pi\to K))} > 1040\,\text{MeV}/c^2.$$

$B^+ \to K^+\ell^+\ell^-$ background

The $B^+ \to K^+ \ell^+\ell^-$ decay can be a source of background if a soft pion from the rest of the event is associated to the kaon or if the kaon is misidentified as pion and a soft kaon from the rest of the event is associated to it. In both cases, a fake K^{*0} candidate is reconstructed. The invariant mass of such B^0 candidates peaks at values above 5380 MeV/c^2 and hence contaminates the upper B^0 invariant mass sideband used in the multivariate analysis described in Section 6.7.5. The contamination from this decay is reduced by requiring the maximum between the three-body invariant masses $m_{(\ell\ell K)}$ and $m_{(\ell\ell(\pi\to K))}$ to be smaller than 5100 MeV/c^2, that is:

$$\max(m_{(\ell\ell K)}, m_{(\ell\ell(\pi\to K))}) < 5100\,\text{MeV}/c^2.$$

$\Lambda_b^0 \to pK^-J/\psi(\to \ell^+\ell^-)$ and $\Lambda_b^0 \to pK^-\ell^+\ell^-$ backgrounds

The $\Lambda_b^0 \to pK^- J/\psi\ (\to \ell^+\ell^-)$ and $\Lambda_b^0 \to pK^-\ \ell^+\ell^-$ decays can be a source of background if the proton is misidentified as pion or if there is a double misidentification in which the proton is misidentified as kaon and the kaon is misidentified as pion. This contribution is reduced by discarding B^0 candidates for which:

$$5575 < m_{(K(\pi\to p)\ell\ell)} < 5665\,\text{MeV}/c^2$$

and $\text{DLL}_{p\pi}$ of the pion is larger than zero (that is, the pion is likely to be a proton), and candidates for which:

$$5575 < m_{((K\to p)(\pi\to K)\ell\ell)} < 5665\,\text{MeV}/c^2$$

and $\text{DLL}_{K\pi}$ of the pion is larger than zero (that is, the pion is likely to be a kaon).

The effectiveness of this veto is studied on the $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$ control channels. The contamination from $\Lambda_b^0 \to pK^-J/\psi\ (\to \ell^+\ell^-)$ decays in the data sample after the full selection is applied is estimated to be at the $0.1 - 0.2\%$ level. The number of expected $\Lambda_b^0 \to pK^-J/\psi\ (\to \ell^+\ell^-)$ decays is evaluated according to

$$\mathcal{N}^{\Lambda_b^0}_{\text{data}} = \mathcal{N}^{B^0}_{\text{data}} \frac{f_{\Lambda_b^0}}{f_d} \frac{\mathcal{B}(\Lambda_b^0 \to pK^-J/\psi)}{\mathcal{B}(B^0 \to K^{*0}J/\psi)} \frac{\epsilon^{\Lambda_b^0}_{\text{tot}}}{\epsilon^{B^0}_{\text{tot}}},$$

where $\mathcal{N}^{B^0}_{\text{data}}$ is the yield of selected $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$ data candidates in the B^0 invariant mass window between 4500 and 5800 MeV/c^2, $f_{\Lambda_b^0}/f_d$ is twice $f_{\Lambda_b^0}/(f_u + f_d)$, with the latter taken from Ref. [130], and ϵ_{tot} is the reconstruction, trigger, and selection efficiency obtained from simulation. The geometric acceptance, whose values are shown in Table 6.4, is also taken into account in the efficiency calculation.

$B^0 \to D^-(\to K^{*0}\ell^-\bar{\nu}_\ell)\ell^+\nu_\ell$ background

The $B^0 \to D^-\ (\to K^{*0}\ell^-\overline{\nu}\ _\ell)\ell^+\nu\ _\ell$ decay has a branching ratio four orders of magnitude larger than that of the signal and can be a source of background if the two neutrinos carry low energies.

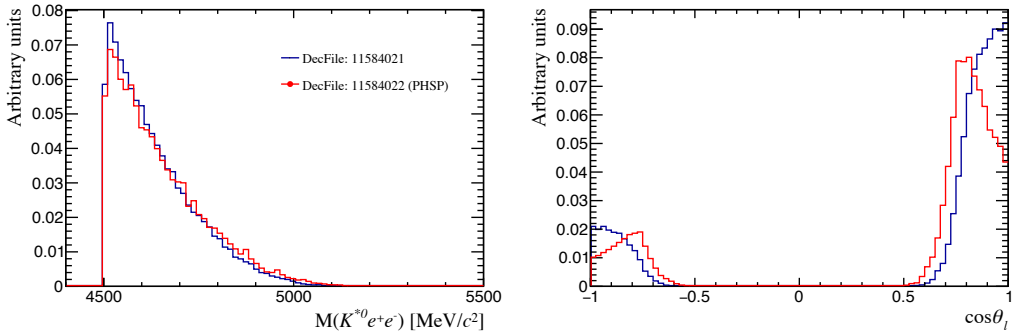


Figure 6.10: Distributions of the $B^0 \to D^- (\to K^{*0}\ell^-\overline{\nu}_\ell)\ell^+\nu_\ell$ simulated candidates. The two colors refer to two different models used in the generation of the decays.

In previous measurements performed at LHCb, two alternative strategies were used to veto this contribution. The first, reported in Ref. [131], required the invariant mass of the $K^{*0}\ell^-$ pair to be larger than 1900 MeV/c^2, which is above the D^- mass. The second, reported in Refs. [34, 39], asked for $|\cos\theta_\ell| < 0.8$. The first strategy is expected to introduce a bias in the angular distribution of the particles in the final state and is hence not suitable for the measurement described in this thesis. The second strategy does not have this problem since the selection in $\cos\theta_\ell$ is symmetric but has the disadvantage of reducing the sensitivity to the angular observables. The rejection power of the second strategy is investigated on two simulation samples of $B^0 \to D^- (\to K^{*0}\ell^-\overline{\nu}_\ell)\ell^+\nu_\ell$ decays, generated according to two different models for the D^- decay. The plots in Fig. 6.10 show the $K^{*0}e^+e^-$ invariant mass and $\cos\theta_\ell$ distributions of the simulated candidates before the reconstruction in the LHCb detector. It is interesting to note that the distributions can be quite different depending on the model used in the generation. The contamination from this background is evaluated by applying the reconstruction and selection requirements to the simulated candidates and by studying how many background decays are expected in the B^0 invariant mass region where the fit is performed.

In the low q^2 bin, the contamination from $B^0 \to D^- (\to K^{*0}\ell^-\overline{\nu}_\ell)\ell^+\nu_\ell$ decays is found to be negligible and, therefore, no requirement on $\cos\theta_\ell$ is applied. In the central q^2 bin, the rejection power of the $\cos\theta_\ell$ requirement is estimated to be between 60% and 80% after the full selection is applied. Since a non-negligible contamination from this background remains even if the $\cos\theta_\ell$ requirement is applied and given the fact that this background has an exponential distribution in the B^0 invariant mass and can therefore be accommodated in the fit, it is preferred to not apply the $\cos\theta_\ell$ requirement to this measurement. It is important to note that this is feasible in the counting method approach but not in the full angular maximum likelihood fit, where the $\cos\theta_\ell$ distribution of this background would have to be taken into account.

Hadron-lepton and hadron-hadron swaps

The hadron-lepton swaps are rejected by requiring that the invariant mass of the hadron-lepton candidate is not compatible with the J/ψ and $\psi(2S)$ resonances if the mass of the hadron is replaced by the mass of the lepton:

$$|m_{((h\to\ell)\ell)} - m_{J/\psi}| > 60\,\text{MeV}/c^2$$

and

$$|m_{((h\to\ell)\ell)} - m_{\psi(2S)}| > 60\,\text{MeV}/c^2.$$

The hadron-hadron swaps are instead vetoed by requiring the $\text{DLL}_{K\pi}$ of the kaon to be larger than the $\text{DLL}_{K\pi}$ of the pion, that is, by requiring that the kaon candidate is more likely to be a kaon than the pion candidate.

Partially reconstructed decays

Partially reconstructed decays are decays for which at least one particle in the final state has not been reconstructed by the detector. These decays can contribute to the background if the reconstructed particles in the final state are the same of the signal and if the resulting invariant mass is within the region considered in the fit. In this measurement, the dominant contribution to this background originates from $B^+ \to K^{*(**)}\ell^+\ell^-$ and $B^+ \to K^{*(**)}J/\psi\,(\to \ell^+\ell^-)$ decays, where $K^{*(**)}$ represents a generic kaon resonance heavier than K^{*0}. Typical decays of such resonances consist of a $K\pi$ pair together with one or more pions and can pass the selection requirements described in the previous sections if one of the pions in the final state is not reconstructed. If the decay is reconstructed as $B^0 \to K^{*0}\ \ell^+\ell^-$ or $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$, the resulting B^0 invariant mass is smaller than the nominal value and is expected to peak at nearly $5140\,\text{MeV}/c^2$ if one pion is missing.

For the $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$ decay mode, this background contribution can be easily vetoed by requiring the PV- and J/ψ-constrained invariant mass of the B^0 candidates to be larger than $5150\,\text{MeV}/c^2$. The residual contamination is estimated from $B^+ \to K_1^+ J/\psi$ simulated candidates, with $K_1^+ \to K^+\pi^+\pi^-$ and $J/\psi \to e^+e^-$. The ratio f_u/f_d is approximated to unity. The residual contamination is found to be below the percent level and can hence be neglected in the model used in the fit. For the $B^0 \to K^{*0}\mu^+\mu^-$ decay mode, partially reconstructed decays are vetoed by requiring the PV-constrained B^0 invariant mass of the signal candidates to be larger than $5150\,\text{MeV}/c^2$. This is not feasible for the $B^0 \to K^{*0}e^+e^-$ decay mode, so alternative approaches are investigated and discussed in detail in Section 6.7.6.

6.7.3 Calorimeter acceptance

As shown in Ref. [34], the cells in the innermost region of the electromagnetic calorimeter, corresponding to the rectangle with $|x| < 363.6\,\text{mm}$ and $|y| < 282.6\,\text{mm}$, are not read out during the data taking. Since this might cause some differences between data and simulation, the $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}J/\psi\ (\to e^+e^-)$ candidates having at least one lepton track traversing this central region are discarded from both data and simulation samples. The distribution on the (x, y) plane of the electromagnetic calorimeter corresponding to the lepton tracks associated to the selected $B^0 \to K^{*0}e^+e^-$ data candidates is shown in Fig. 6.11.

6.7.4 Charmonium resonances

Charmonium resonances, such as the $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$ decay, are a peaking background for the $B^0 \to K^{*0}\ \ell^+\ell^-$ decay in the central q^2 region. This background is negligible for decays to muons but not for decays to electrons, where the lower resolution and the effects induced by bremsstrahlung emission are expected to play a role. Therefore, a study of the expected size of this background in the electron mode is needed. A novel stragegy is pursued in this measurement, compared to

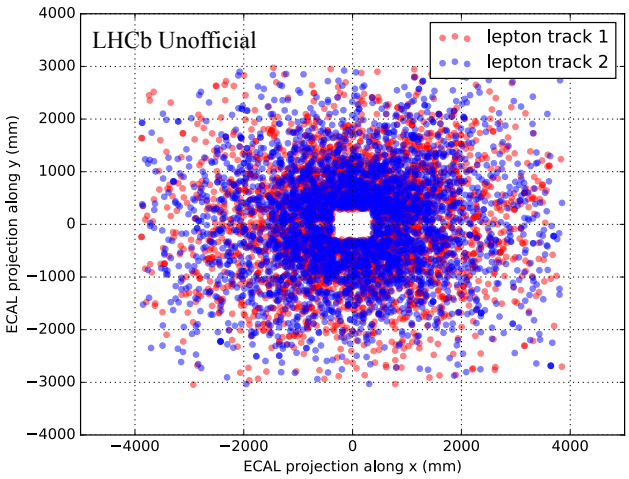


Figure 6.11: Distribution on the (x, y) plane of the electromagnetic calorimeter corresponding to the lepton tracks associated to the $B^0 \to K^{*0} e^+ e^-$ data candidates selected by the calorimeter acceptance requirement. The two leptons in the final state are highlighted in red and blue.

previous measurements at LHCb. The contamination from $B^0 \to K^{*0} J/\psi$ $(\to e^+ e^-)$ decays leaking into the central q^2 region is suppressed by selecting not only on the q^2, but also on a novel variable, the constrained q^2. This can provide a significant increase in the statistics for the rare mode, as discussed in the following.

The efficiency of the combined selection is investigated on simulation and data samples of $B^0 \to K^{*0} J/\psi$ $(\to e^+ e^-)$ candidates, as well as on the simulation sample of the rare mode $B^0 \to K^{*0} e^+ e^-$. The gain in selecting on both normal and constrained q^2 can be seen in Fig. 6.12. The plots show the residual background due to the leakage of the radiative tail of the J/ψ in the rare mode for different selections of q^2, separately for each trigger category. By selecting also on the constrained q^2 it is possible to have a larger bin in the central q^2 region, up to $7.0\,\mathrm{GeV}^2/c^4$ instead of $6.0\,\mathrm{GeV}^2/c^4$, with even a smaller background contamination than when considering a maximum q^2 of $6.0\,\mathrm{GeV}^2/c^4$.

The distribution of the $B^0 \to K^{*0}\, \ell^+ \ell^-$ and $B^0 \to K^{*0} J/\psi$ $(\to \ell^+ \ell^-)$ data candidates as a function of PV-constrained B^0 invariant mass $m(K^+ \pi^- \ell^+ \ell^-)_{\mathrm{PV}}$ and q^2 is shown in Fig. 6.13 for both the muon mode (left) and the electron mode (right). The analogous distribution as a function of the constrained q^2 is shown in Fig. 6.14. The two horizontal bands in Fig. 6.13 corrresponds to the J/ψ and $\psi(2S)$ resonances, while the diagonal bands correspond to the radiative tails generated by bremsstrahlung emission and are, as expected, more pronounced in the electron mode. Due to the constraint on q^2, the $B^0 \to K^{*0}\, \ell^+ \ell^-$ and $B^0 \to K^{*0} J/\psi$ $(\to \ell^+ \ell^-)$ data candidates appear rotated in Fig. 6.14. For the electron mode, only the $B^0 \to K^{*0} J/\psi$ $(\to \ell^+ \ell^-)$ data candidates are visible, while both $B^0 \to K^{*0} J/\psi$ $(\to \ell^+ \ell^-)$ and $B^0 \to K^{*0}\, \ell^+ \ell^-$ data candidates are visible in the case of the muon mode.

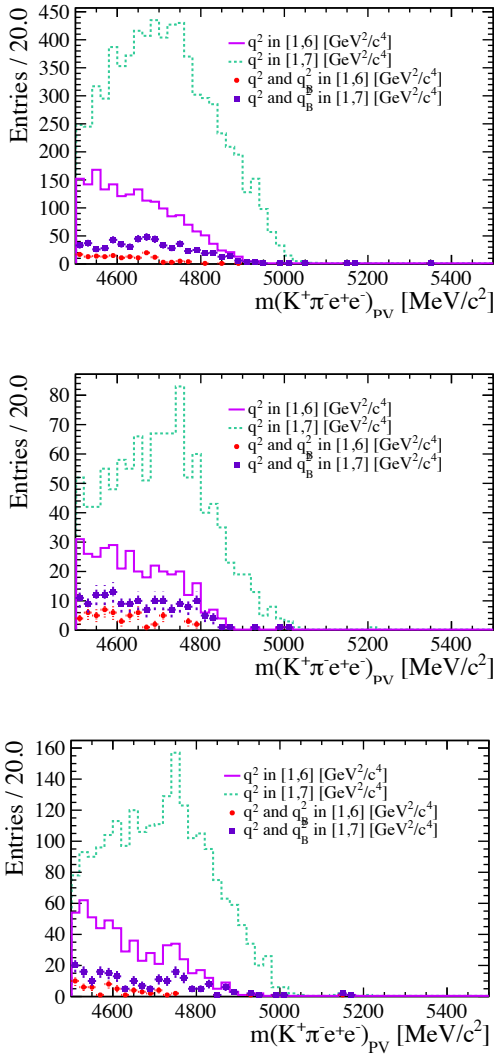


Figure 6.12: Distribution of the PV-constrained B^0 invariant mass of the background due to $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) candidates leaking in the central q^2 bin for the `L0E` (top), `L0H` (middle), and `L0TIS` (bottom) trigger categories, as obtained from simulation.

6.7.5 Multivariate classifier

Random combinations of particles from the same pp collision can be a source of background if they have invariant masses compatible with those of the signal candidates. A large fraction of this *combinatorial background* is removed by means of a multivariate classifier. To this purpose, machine learning techniques implemented in the `scikit-learn` python package [1] and in the Reproducible Experiment Platform (`REP`) [2, 3] and High-Energy-Physics Machine Learning (`hep_ml`) [4] python libraries are used. A multivariate classifier is first trained and tested on two samples labelled as signal and background and then applied to the samples from which the combinatorial background has to be rejected. The classifier response is set to be between zero and unity, with values close to zero (unity) indicating a large probability of the reconstructed candidate to be background (signal). An optimisation procedure allows to determine the threshold on the classifier response that guarantees

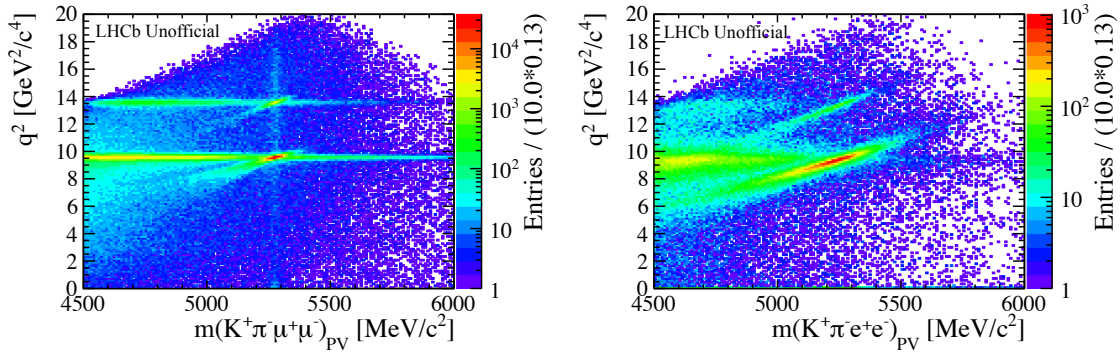


Figure 6.13: Distribution of the $B^0 \to K^{*0}\ \ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$ data candidates as a function of PV-constrained B^0 invariant mass $m(K^+\pi^-\ell^+\ell^-)_{\rm PV}$ and q^2, as obtained for the muon mode (left) and the electron mode (right).

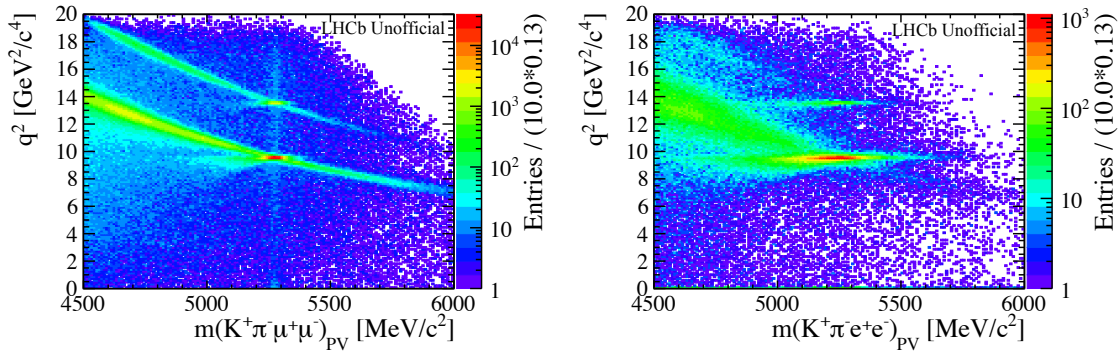


Figure 6.14: Distribution of the $B^0 \to K^{*0}\ \ell^+\ell^-$ and $B^0 \to K^{*0}J/\psi\ (\to \ell^+\ell^-)$ data candidates as a function of PV-constrained B^0 invariant mass $m(K^+\pi^-\ell^+\ell^-)_{\rm PV}$ and constrained q^2, as obtained for the muon mode (left) and the electron mode (right).

the best compromise between *signal efficiency* and *background rejection*.

Signal and background samples

Two different classifiers are used to select signal decays to electrons and muons. Training and testing are performed on combined samples of 2011 and 2012 and, for electrons, jointly for the three trigger categories. This choice is driven by the observation that the statistics available is the limiting factor in the classifier performance. An increase in the statistics is hence more beneficial than the use of samples recorded under the same conditions (mainly the energy of the pp collisions) and with the same kinematic properties. On the other hand, the optimisation of the classifiers is performed separately for each trigger category. The low and central q^2 bins are combined for the training but considered separately in the optimisation.

The signal is obtained from $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ simulation samples that are corrected to account for differences between simulation and data, while the B^0 candidates populating the upper sideband of the B^0 invariant mass distribution in data are used as background sample.

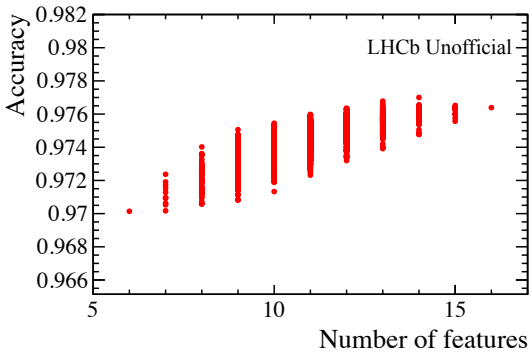


Figure 6.15: Accuracy of each combination of variables as a function of the number of variables in the combination. Each red dot in the plot corresponds to a specific combination.

For electrons, only candidates with $m_{B^0} > 5600\,\text{MeV}/c^2$ are used in the background sample. For muons, this requirement is loosened to $m_{B^0} > 5400\,\text{MeV}/c^2$. The tighter selection of the electron mode is due to the worse resolution and to the fact that the mass region between $5400\,\text{MeV}/c^2$ and $5600\,\text{MeV}/c^2$ is populated by decay candidates with a bremsstrahlung overcorrection. Signal and background candidates must satisfy the same requirements described in Section 6.7.1 apart from the requirement on the $K\pi$ invariant mass. This requirement is in fact loosened to $200\,\text{MeV}/c^2$ around the nominal value of the K^{*0} mass to increase the statistics available for the training and testing.

Selection of discriminating variables

The variables that guarantee the best separation between signal and background are obtained from the electron samples by considering all possible combinations of at least six of the following variables:

- p_T, impact parameter χ^2, flight distance χ^2, decay vertex χ^2/ndf, `DIRA`, and χ^2 of the kinematic fit of the B^0 candidate;
- decay vertex χ^2/ndf of the K^{*0} candidate;
- decay vertex χ^2/ndf of the dilepton pair;
- sum of the p_T and sum of the impact parameter χ^2 of the hadrons (leptons) in the final state;
- minimum and maximum p_T and impact parameter χ^2 of the hadrons (leptons) in the final state;
- ProbNN variables of the four particles in the final state.

The `DIRA` variable corresponds to the cosine of the angle between the momentum of the B^0 candidate and the vector from the PV to the decay vertex of the B^0 candidate. The variables related to the B^0 candidate are included in all combinations since they are expected to provide the best discrimination between signal and background. A *gradient boosting* classifier is trained using as input the variables in each combination. The accuracy of the classifier, which corresponds to the fraction of candidates in the testing sample that are labelled correctly as signal or background, is then used to quantity

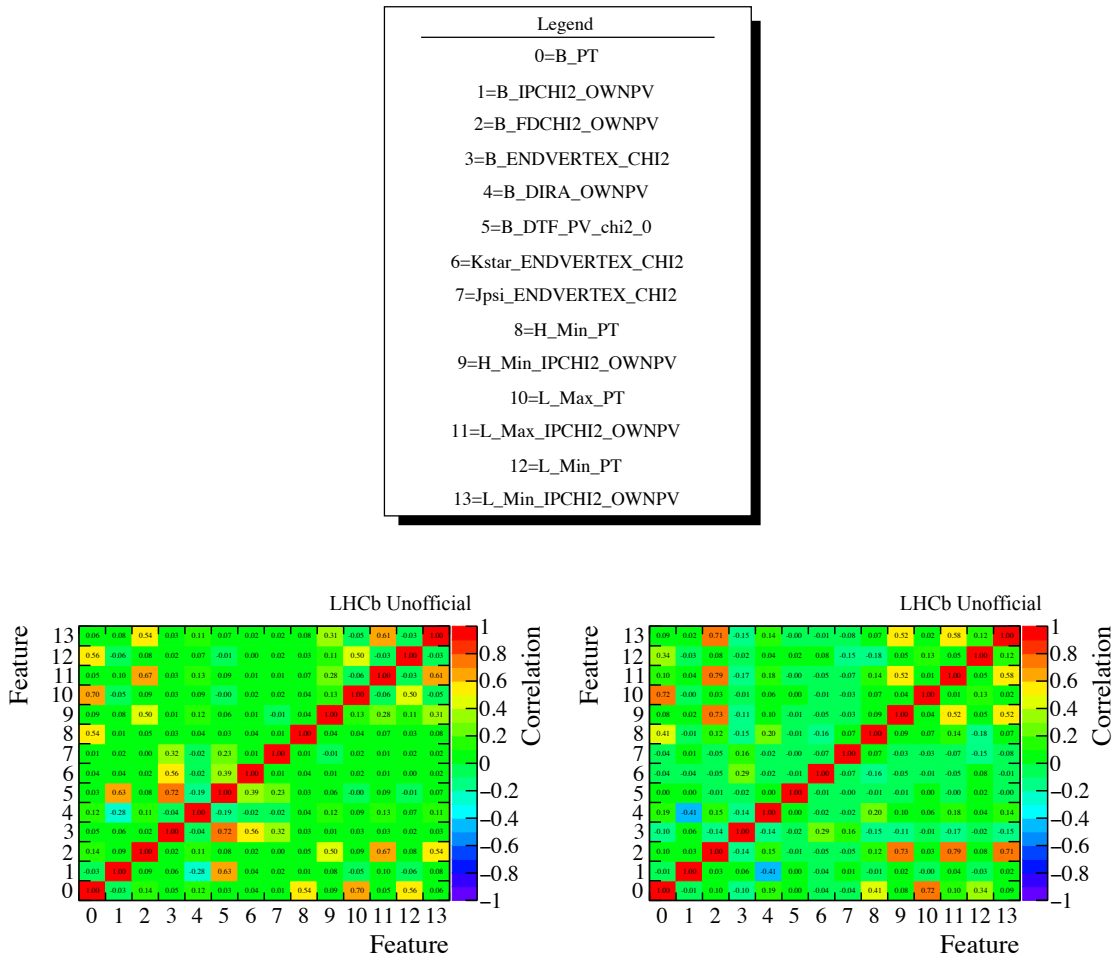


Figure 6.16: Correlation among the best discriminating variables for the signal (left) and background (right) samples. The numbers on the axes refer to the variables reported in the legend.

the discriminating power of a given combination. The optimal set of variables to be used in the classification is given by the combination associated to the highest accuracy.

Some considerations on this initial set of variables can be made. The PID variables do not seem to provide a good discriminating power. Therefore, they are removed from this procedure in the first iteration. The sum of the p_{T} and the sum of the impact parameter χ^2 are highly correlated with the minimum and maximum p_{T} and impact parameter χ^2 and provide a worse discrimination, so they are also ignored in the following. The results for the iterative procedure on the remaining variables are summarised in Fig. 6.15, which shows the accuracy of each combination as a function of the number of variables in the combination. The accuracy does not have a strong dependence on the number of variables taken into account, though a gradual improvement is visible when increasing the number of variables from six to 14. The accuracy is maximum when 14 variables are used in the training and decreases slightly if additional degrees of freedom have to be taken into account. This is probably due to overfitting, that is, to the fact that, with the statistics available, the classifier is not able to extract useful information when too many variables are provided as input.

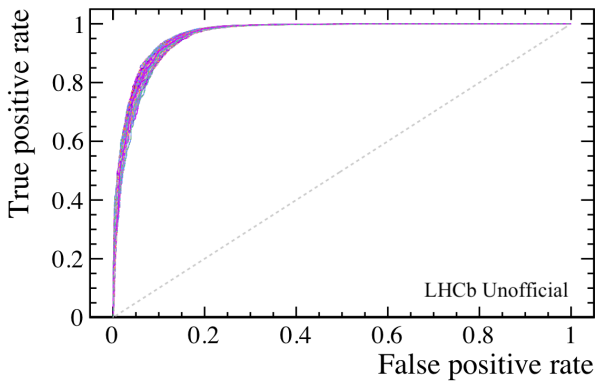


Figure 6.17: ROC curve of each configuration of the `XGBoost` classifier. Each color refers to a specific configuration of the classifier. In total, 432 different configurations of `XGBoost` are tested.

For a given number of variables, the best accuracy is reached when including all B^0 variables. These variables alone are able to provide almost the same discriminating power as the full set of variables described above. The 14 variables resulting in the maximum accuracy are used to discriminate between signal and background. These are the six variables related to the B^0 candidate, the decay vertex χ^2/ndf of the K^{*0} candidate, the decay vertex χ^2/ndf of the dilepton pair, the minimum p_{T} and impact parameter χ^2 of the hadrons, and the minimum and maximum p_{T} and impact parameter χ^2 of the leptons. Their correlation matrices, as obtained from the signal and background samples, are shown in Fig. 6.16. As expected for the signal, the χ^2 of the kinematic fit of the B^0 candidate is highly correlated with the impact parameter χ^2 (63%) and decay vertex χ^2/ndf (72%) of the candidate itself. This does not hold for the background, since the latter consists of random combinations of particles. The K^{*0} and B^0 decay vertex χ^2/ndf are also correlated with each other (56%) in the signal sample, while no correlation appears in the background sample. The p_{T} of the B^0 candidate is highly correlated with the minimum and maximum p_{T} of the leptons (56% and 70%, respectively) and, to a lesser extent, to the minimum p_{T} of the hadrons (54%). Two other pairs of highly correlated variables are given by the flight distance χ^2 of the B^0 candidate and the maximum (67%) and minimum (54%) impact parameter χ^2 of the leptons. Similar correlations are observed in the muon mode.

Selection of classifier

After choosing the optimal set of variables, several classifiers are optimised and then tested against each other. The following classifiers are considered:

- `TMVA`;
- `GradientBoosting`;
- `XGBoost`;
- `DTC` (Decay Tree Classifier);
- `EXC` (Extra Trees Classifier).

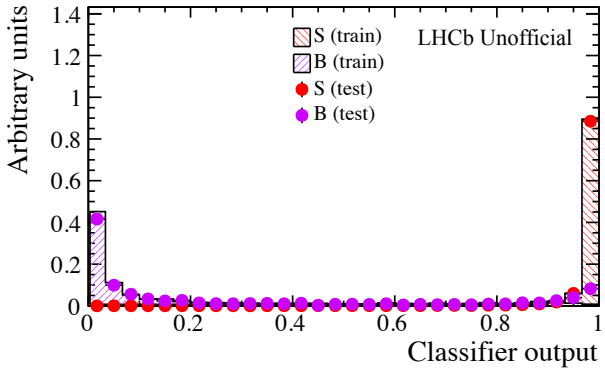


Figure 6.18: Comparison among the classifier responses to the training and testing samples of signal and background, corresponding to the selected configuration of `XGBoost`.

The optimisation of a given classifier is achieved by performing a grid search in its parameter space, that is, by evaluating the performance of the classifier for different settings of the available parameters.

Different configurations of the same classifier are compared to each other by plotting their *Receiver Operating Characteristic* (ROC) curves. These show the dependence between *true positive rate*, corresponding to the fraction of signal candidates that are correctly labelled, and *false positive rate*, corresponding to the fraction of background candidates that are wrongly labelled, for different thresholds of the classifier response. The ideal classifier would have true positive rate equal to unity and false positive rate equal to zero. In reality, the true positive rate is equal to unity only for a threshold equal to zero, which corresponds to the maximum false positive rate, and goes from unity to zero as the threshold increases. An example is shown in Fig. 6.17. The information carried by the ROC curve can be condensed in one number by using the *Receiver Operating Characteristic Area Under the Curve* (ROC AUC), which, as the name suggests, corresponds to the integral of the ROC curve. The configuration associated to the largest ROC AUC is chosen as the optimal configuration of the classifier. The improvement achievable by tuning the configuration of each classifier is shown in Fig. 6.17 for `XGBoost`.

The `GradientBoosting` and `XGBoost` algorithms provide the highest ROC AUC, with `XGBoost` performing slightly better than `GradientBoosting`. The configuration corresponding to the best ROC AUC of `XGBoost` is then used as the nominal multivariate classifier for this measurement. The performance of this configuration is summarised in Figs. 6.18, 6.19, and 6.20. In particular, the classifier response on the training and testing samples of signal and background is shown in Fig. 6.18, the feature importance is shown in Fig. 6.19, and the efficiency as a function of the B^0 invariant mass and of the decay angles ϕ, $\cos\theta_K$, and $\cos\theta_\ell$ is shown in Fig. 6.20 for both signal and background samples and for several thresholds of the classifier response. The classifier is flat in the three decay angles, but shows a linear dependence of the efficiency on the B^0 invariant mass for both signal and background.

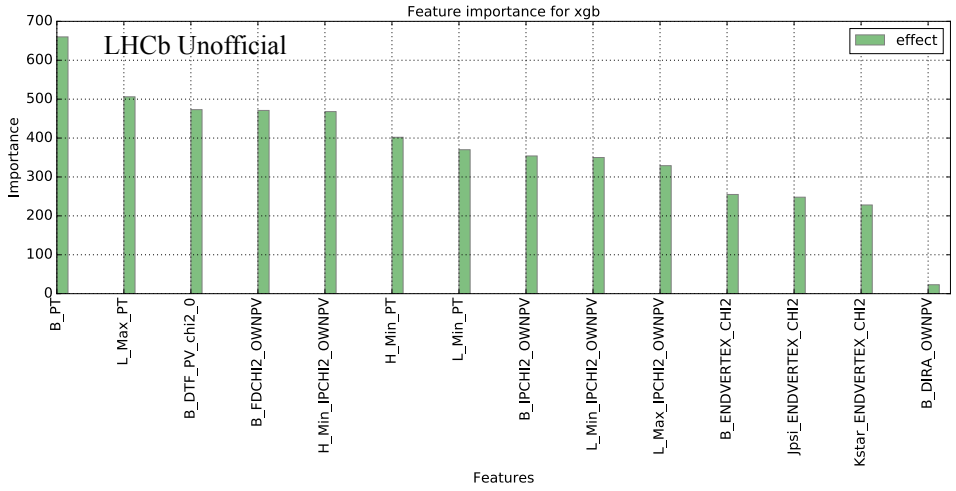


Figure 6.19: Feature importance, corresponding to the selected configuration of `XGBoost`.

Uniform classifiers

The possibility to use a uniform classifier to discriminate between signal and background is also investigated. A uniform classifier can be trained in a way that guarantees a uniform classifier response over one or more variables. Such uniformity, in particular, can be required for either the signal or background sample used in the training. One typical situation in which uniformity in the predictions can be useful is when no prior information on the dynamics of a given phase-space is available. In this case, a uniform classifier reduces potential risks of biasing the signal distribution of interest. A typical example in particle physics is the search for a new particle. In this case, a classifier with uniformity in the background predictions over the invariant mass of the signal candidates reduces the risk to create a peaking background. In this measurement, the invariant mass of the B^0 candidates and the decay angles θ_ℓ, θ_K, and ϕ are the variables of interest. The S_i observables are in fact computed as the asymmetry between two categories of signal candidates, S_i^+ and S_i^-, whose yields are obtained from a fit to the invariant mass of the B^0 candidates, separated according to the values of the decay angles θ_ℓ, θ_K, and ϕ. If the classifier response is not uniform in these variables, a bias might appear in the measured asymmetry.

The performance of the `XGBoost` classifier obtained from the optimisation described above is compared to that achievable by using the following uniform classifiers:

- `uBoost`;
- `UGBknnAda`;
- `UGBFlatnessLoss`.

For each classifier listed above, two configurations are tested, one with uniformity in the B^0 invariant mass only and one with uniformity also in the decay angles θ_ℓ, θ_K, and ϕ. The second configuration is expected to provide lower discriminating power, since more constraints have to be accommodated. It is found, however, that the performance of the uniform classifiers does not change significantly when imposing uniformity over additional variables. Additional information on the performance of the uniform classifiers can be found in Appendix A.

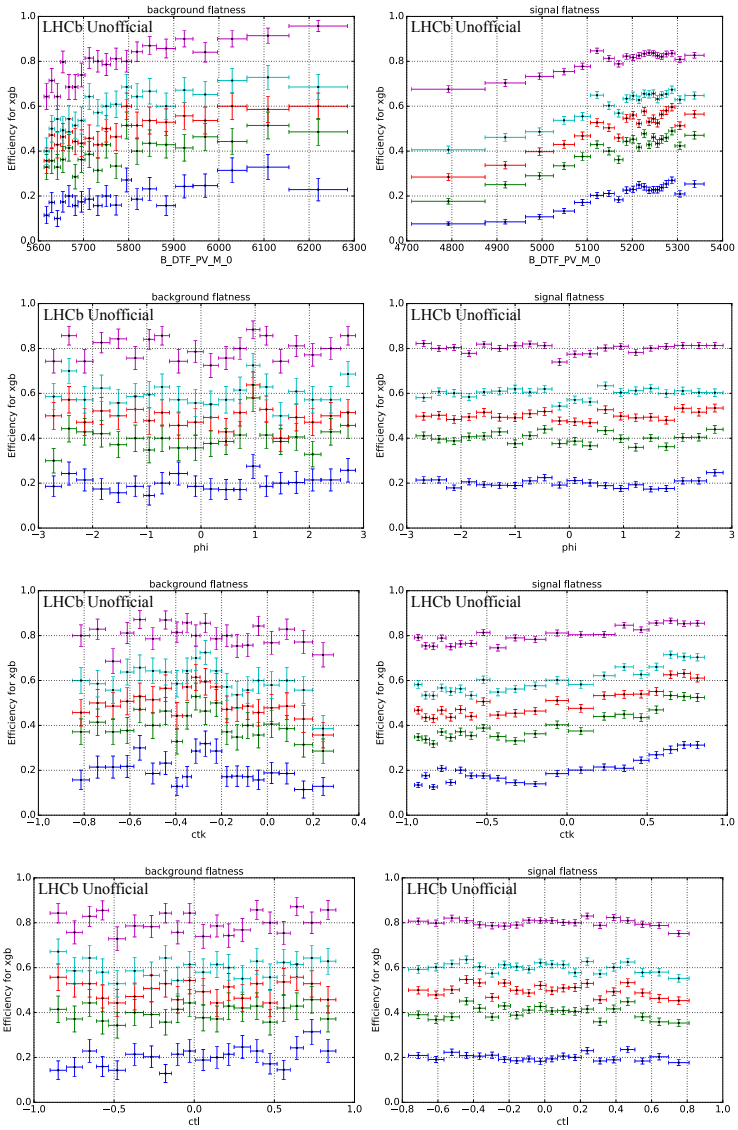


Figure 6.20: Efficiency as a function of the B^0 invariant mass (first row) and of the decay angles ϕ (second row), $\cos\theta_K$ (third row), and $\cos\theta_\ell$ (fourth row) for the background (left) and signal (right) samples, corresponding to the selected configuration of `XGBoost`. The different colors refer to several thresholds of the classifier response and are chosen to cover uniformly the full efficiency range from zero to unity.

Since the overall performance of the uniform classifiers is worse than that achievable using `XGBoost`, and given that the improvement in uniformity with respect to `XGBoost` is negligible, it is preferred to use `XGBoost`, configured as described in the previous paragraphs.

k-folding

In order to make use of the full statistics of the signal and background samples in the training, a k-folding approach is adopted. This strategy consists in splitting the input sample randomly in k subsamples s_i, with $i \in \{1...k\}$. After that, k different classifiers are trained. The classifier j is trained on the union of the s_i subsamples, with $i \neq j$, and then tested on the excluded subsample s_j.

This procedure is repeated k times, so that all combinations of k−1 subsamples out of k are used in the training phase. The k classifiers are then combined together and used to classify candidates that have not been used during the training and testing. This can be achieved by selecting one classifier randomly and using its prediction or by averaging the predictions of all classifiers. For the measurement described here, the former method has been preferred.

As k increases, more and more candidates are used in the training of each classifier. This is expected to increase the stability of the classifier response and decrease the risk of overfitting. On the other hand, the testing sample becomes smaller with increasing k, so the uncertainty on the performance assessment increases. In addition, given two classifiers c_j and c_k, the overlap between the candidates used in the training of c_j and in the training of c_k increases with k. At the extreme case in which all but one candidate out of N are used to train each of N classifiers, the overlap is maximum. In order to find the optimal value of k, three different possibilities are tested: k= 2, k= 5, and k= 10. No difference in the results is observed when varying the value of k. The folding with k= 10 is chosen as default strategy since it is consistent with previous measurements at LHCb.

Optimisation of multivariate classifier threshold

In order to lower the contamination from combinatorial background, only signal candidates with classifier response larger than a given threshold are retained. The value of the threshold is optimised separately for electron and muon modes and for the low and central q^2 bins. For the electron mode, in addition, the threshold is optimised independently for each trigger category.

The optimisation is performed by using the figure of merit $S/\sqrt{S+B}$, with S and B being the signal and background yields, respectively.

The signal yield S is calculated as

$$\mathcal{N}_S = \mathcal{N}_{J/\psi} \frac{\mathcal{B}_S}{\mathcal{B}_{J/\psi}} \frac{\epsilon_S}{\epsilon_{J/\psi}}, \tag{6.3}$$

where $\mathcal{N}_{J/\psi}$ is the yield of $B^0 \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ obtained from a fit to the data, $\mathcal{B}_S$ and $\mathcal{B}_{J/\psi}$ are the braching fractions of the $B^0 \to K^{*0}\ \ell^+\ell^-$ and $B^0 \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ decays, and ϵ_S and $\epsilon_{J/\psi}$ are their total efficiencies. In order to compute $\mathcal{N}_{J/\psi}$ for each value of the threshold, the PV- and J/ψ-constrained invariant mass distribution of the $B^0 \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ candidates is fitted similarly to what described in Section 6.6.2. In doing so, the $B^0 \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ simulated and data candidates passing the selection described in Sections 6.5, 6.7.1, 6.7.2, and 6.7.3 are required to have a q^2 and a constrained q^2 between 7.0 and 11.0 GeV^2/c^4. Only candidates with classifier response larger than the threshold are taken into account. The results of the fit to the simulation sample are used to partially constrain the model used to fit the data sample. The models used in the fit are the same described in Section 6.6.2, except for an additional constraint on the yield of the $B^0_s \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ component when fitting the data sample. In order to improve the stability of the fit, in fact, the yield of the $B^0_s \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ component is constrained to that of the $B^0 \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ scaled by the fragmentation fraction ratio f_s/f_d and the branching fraction ratio $\mathcal{B}\,(B^0_s \to J/\psi\, K^{*0},\ J/\psi \to e^+e^-)/\mathcal{B}\,(B^0 \to J/\psi\, K^{*0},\ J/\psi \to e^+e^-)$, as reported in Ref. [132] and in Ref. [15], respectively. The fit to $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ simulated and data candidates belonging to the `L0E` trigger category prior to any cut on the classifier response is shown in Fig. 6.21 as a reference.

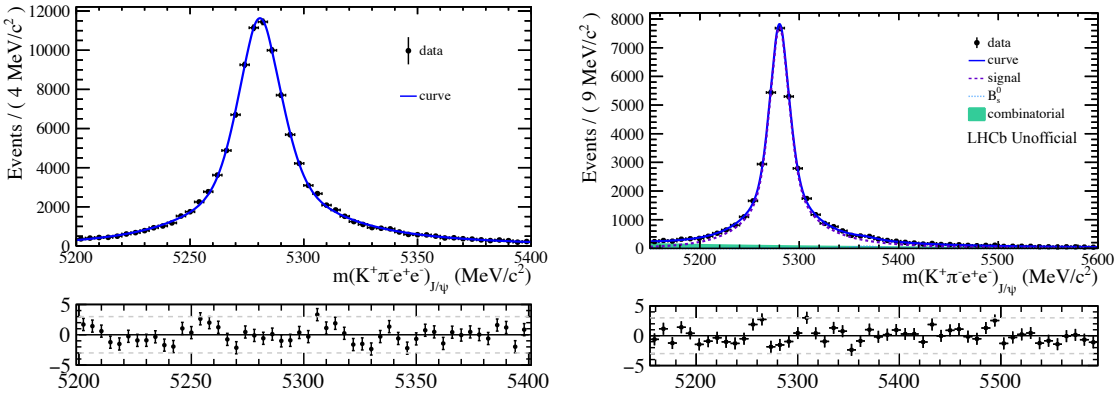


Figure 6.21: Fit to the PV and J/ψ-constrained B^0 invariant mass distribution of the $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) simulated (left) and data (right) candidates belonging to the `L0E` trigger category prior to any cut on the classifier response.

The branching fraction of the $B^0 \to K^{*0} J/\psi$ ($\to \ell^+\ell^-$) decay, which corresponds to $\mathcal{B}_{J/\psi}$ in Eq. (6.3), is set to

$$\mathcal{B}(B^0 \to J/\psi K^{*0}, J/\psi \to \ell^+\ell^-) = (7.6 \pm 0.3) \times 10^{-5},$$

where the values reported in Ref. [15] are used. For the rare mode, the branching fraction, which corresponds to $\mathcal{B}_S$ in Eq. (6.3), yields

$$\mathcal{B}(B^0 \to K^{*0} e^+e^-) = (1.03 \pm 0.19) \times 10^{-6},$$
$$\mathcal{B}(B^0 \to K^{*0} \mu^+\mu^-) = (1.03 \pm 0.06) \times 10^{-6}.$$

The efficiencies ϵ_S and $\epsilon_{J/\psi}$ in Eq. (6.3) are calculated from the $B^0 \to K^{*0}\ \ell^+\ell^-$ and $B^0 \to K^{*0} J/\psi$ ($\to \ell^+\ell^-$) simulation samples as described in Section 6.8.

The background yield B is obtained by counting the number of $B^0 \to K^{*0}\ \ell^+\ell^-$ data candidates with PV-constrained B^0 invariant mass larger than 5600 MeV/c^2. The number of $B^0 \to K^{*0}\ \ell^+\ell^-$ simulated candidates above this value is negligible.

The thresholds that are chosen are summarised in Table 6.3, for each decay mode, bin of q^2, and trigger category. The figure of merit obtained for the muon mode in the low and central q^2 bins is shown as reference in Fig. 6.22, while the corresponding plots for the electron mode and, in particular, for the `L0E` trigger category, are given in Fig. 6.23. The signal and background yields as a function of the threshold on the classifier response for the same samples are shown in Figs. 6.24 and 6.25, respectively.

A series of cross checks is performed to determine the robustness of the figure of merit. The optimal cut on the classifier response is determined using a factor of two or a factor of one half to scale the background yield B. In addition, the following figures of merit are evaluated:

- S/B;
- $S/(S+B)$;
- $S^2/(S+B)^{\frac{3}{2}}$.

All these strategies provide a similar response, which results in a further validation of the nominal

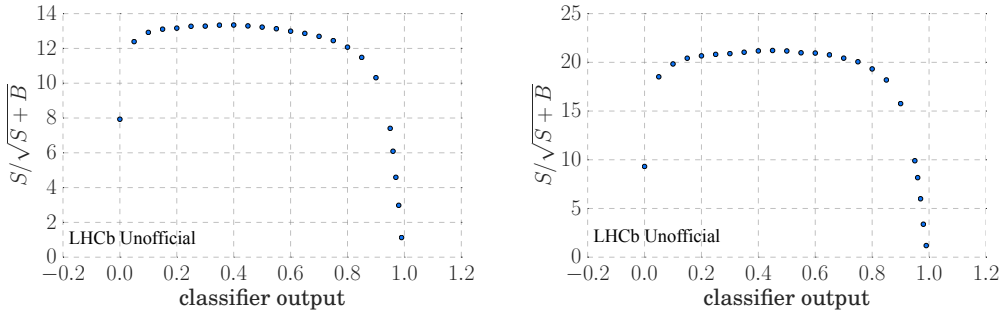


Figure 6.22: Figure of merit for the muon mode in the low (left) and central (right) q^2 bins.

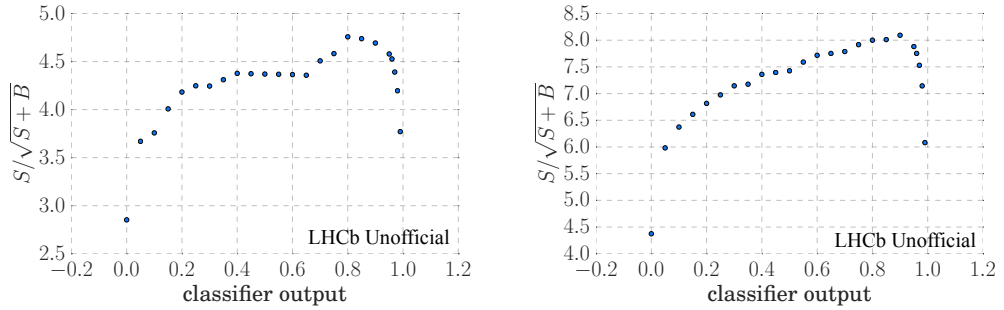


Figure 6.23: Figure of merit for the `L0E` trigger category of the electron mode in the low (left) and central (right) q^2 bins.

approach.

6.7.6 Optimised selection against partially reconstructed decays

Partially reconstructed decays of the type $B^+ \to K^{*(**)}\ell^+\ell^-$ and $B^+ \to K^{*(**)}J/\psi(\to \ell^+\ell^-)$ are a background contribution to the $B^0 \to K^{*0}e^+e^-$ decay, due to the limited mass resolution of the latter. This background contribution is usually modelled in the mass fit (see Section 7.1). In the measurement described in this thesis, however, a series of studies is performed in order to possibly reduce the contamination from such background contribution. Two approaches are investigated: a selection on the so-called HOP variable, aligned to Ref. [34], and a dedicated multivariate classifier. The performance of the two methods is studied on simulated candidates of $B^+ \to K_1^+e^+e^-$ and $B^+ \to K_2^*(1430)^+e^+e^-$. The same-sign $B^0 \to K^{*0}\mu^+\mu^+$ and $B^0 \to K^{*0}\mu^+e^+$ data samples are also used for cross-checks. The $B^0 \to K^{*0}\mu^+\mu^+$ sample is used instead of the $B^0 \to K^{*0}e^+e^+$ sample, since the latter was not available when performing the study.

The distribution of the $B^+ \to K_1^+e^+e^-$ simulated candidates as a function of the PV-constrained B^0 invariant mass $m(K^+\pi^-e^+e^-)_{\rm PV}$ and q^2 is shown in Fig. 6.26 (left), while the analogous distribution as a function of the PV-constrained B^0 invariant mass $m(K^+\pi^-e^+e^-)_{\rm PV}$ and constrained q^2 is shown in Fig. 6.26 (right). These candidates are distributed across the whole q^2 region where the analysis is performed and, although they are more abundant in the low mass region, can have

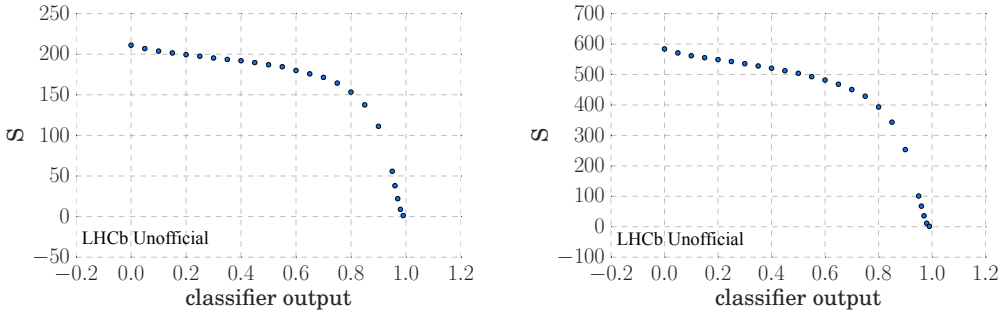


Figure 6.24: Signal yield as a function of the threshold on the classifier response for the muon mode in the low (left) and central (right) q^2 bins.

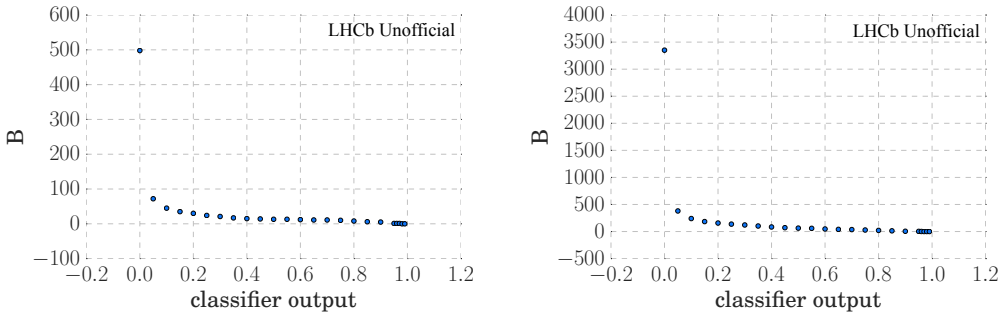


Figure 6.25: Background yield as a function of the threshold on the classifier response for the muon mode in the low (left) and central (right) q^2 bins.

B^0 invariant mass compatible with that of the signal.

HOP approach

In the $B^0 \to K^{*0}e^+e^-$ decay, the vectorial sum of the final state particles is not expected to have a component orthogonal to the flight direction of the B^0 meson. This property does not necessarily hold for $B^0 \to K^{*0}e^+e^-$ candidates that originate from partially reconstructed decays or, more in general, from some kind of background. This different behaviour can hence be exploited to improve the rejection power of the selection described above. To this extent, two new variables, $\alpha_{\rm HOP}$ and $m_{\rm corr}$, are defined. The flight direction of the B^0 meson is determined from the position of its primary and decay vertices. The momenta of the K^{*0} meson and of the dilepton pair orthogonal to the flight direction of the B^0 meson are used to compute the correction factor $\alpha_{\rm HOP}$, according to:

$$\alpha_{\rm HOP} = \frac{p_{\rm T}(K^{*0})}{p_{\rm T}(\ell^+\ell^-)}.$$

A sketch of the variables involved in the $\alpha_{\rm HOP}$ definition is shown in Fig. 6.27. For signal candidates, $\alpha_{\rm HOP}$ is expected to be unity, with possible deviations from unity mainly due to energy loss in the dilepton pair. The energy carried by the bremsstrahlung photons can be recovered by multiplying

Table 6.3: Optimal cut on the classifier response, corresponding to the maximum of the figure of merit, separately for each decay mode, q^2 bin, and trigger category.

Decay mode	Bin of q^2	Trigger category	Threshold
muon mode	low	`L0M`	0.40
	central	`L0M`	0.45
electron mode	low	`L0E`, `L0H`, `L0TIS`	0.80
	central	`L0E`, `L0H`, `L0TIS`	0.90

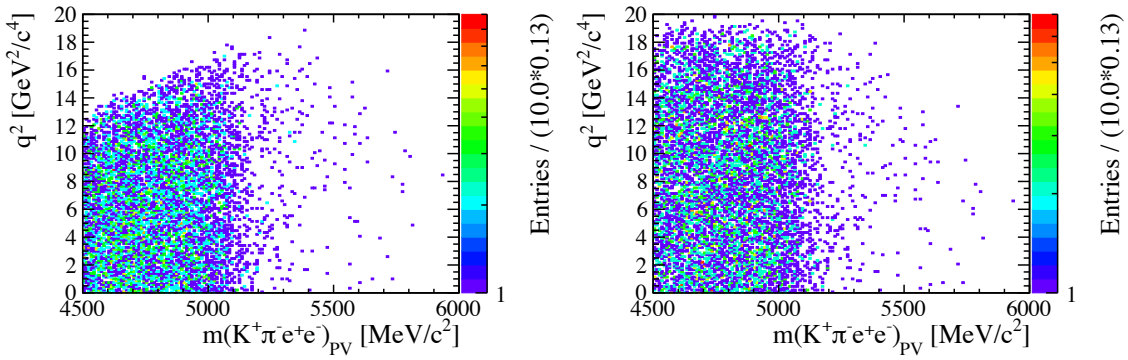


Figure 6.26: Distribution of the $B^+ \to K_1^+ e^+ e^-$ simulated candidates as a function of the PV-constrained B^0 invariant mass $m(K^+\pi^- e^+ e^-)_{\rm PV}$ and q^2 (left) or constrained q^2 (right), after the selection is applied.

the momentum of the dilepton pair orthogonal to the flight direction of the B^0 meson by $\alpha_{\rm HOP}$. Since bremsstrahlung emission is, to a good approximation, collinear to the direction of flight of the lepton, the same correction can be applied to the longitudinal component, giving

$$p_{\rm corr}(\ell^+\ell^-) = \alpha_{\rm HOP}\; p(\ell^+\ell^-).$$

The B^0 invariant mass can then be recalculated by using $p_{\rm corr}(\ell^+\ell^-)$ instead of $p(\ell^+\ell^-)$. According to the definition of $\alpha_{\rm HOP}$, the same correction is applied to both leptons when recalculating the B^0 invariant mass. This approximation neglects the fact that in general the two leptons lose a different amount of energy via bremsstrahlung.

The two-dimensional distribution in $m_{\rm corr}$ and logarithm of the flight distance χ^2 of the $B^0 \to$

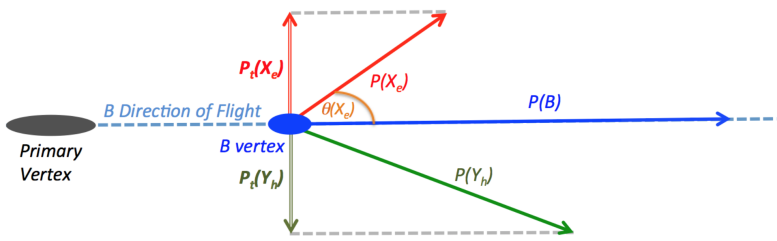


Figure 6.27: Sketch of the variables involved in the $\alpha_{\rm HOP}$ definition.

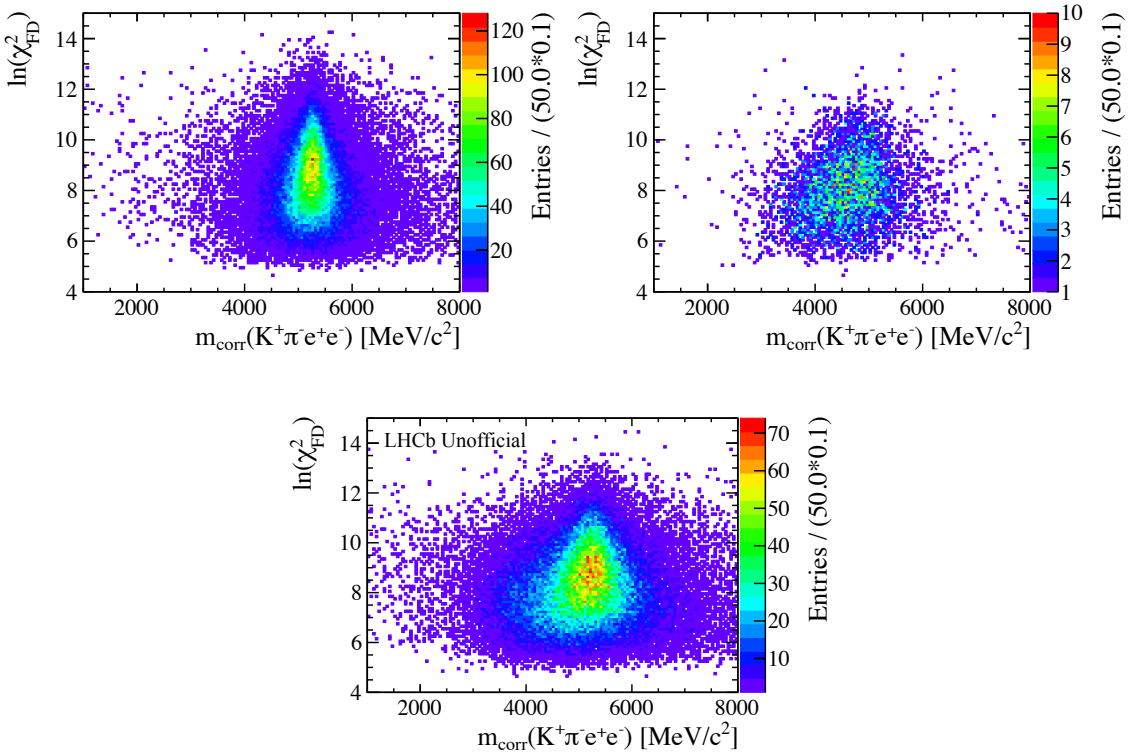


Figure 6.28: Number of $B^0 \to K^{*0}e^+e^-$ (top left) and $B^+ \to K_1^+e^+e^-$ (top right) simulated candidates as a function of $m_{\rm corr}$ and logarithm of the flight distance χ^2, after the selection is applied, compared to the corresponding distribution of the $B^0 \to K^{*0}e^+e^-$ data candidates (bottom). The HOP requirement corresponds to a nearly vertical straight line on the left side of the region where most of the signal is.

$K^{*0}e^+e^-$ and $B^+ \to K_1^+e^+e^-$ simulated candidates after the selection is applied is shown at the top of Fig. 6.28. The analogous distribution for the $B^0 \to K^{*0}e^+e^-$ data candidates is shown at the bottom of Fig. 6.28. The distributions are obtained by combining the low and central q^2 bins. The difference between the distributions of the $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K_1^+e^+e^-$ simulated candidates allows to reduce the contamination from partially reconstructed decays by requiring

$$m_{\rm corr} > 5072 + \ln(\text{FD } \chi^2)$$

for the low q^2 bin and

$$m_{\rm corr} > 4926 + 10\ln(\text{FD } \chi^2)$$

for the central q^2 bin. This selection, which is taken from Ref. [34], has a signal efficiency from 74% to 81% (from 73% to 80%) and a rejection power from 76% to 85% (from 68% to 70%) in the low (central) q^2 bin. The range is due to the different values obtained for the `L0E`, `L0H`, and `L0TIS` trigger categories. The signal efficiency and rejection power are evaluated after the full selection is applied to the samples.

The invariant mass distribution of the candidates selected or discarded by the HOP requirement is shown in Fig. 6.29 for the $B^0 \to K^{*0}e^+e^-$ and $B^+ \to K_1^+e^+e^-$ simulated candidates in the low and in the central q^2 bins. The plots refer to simulated candidates in the `L0E` trigger category only,

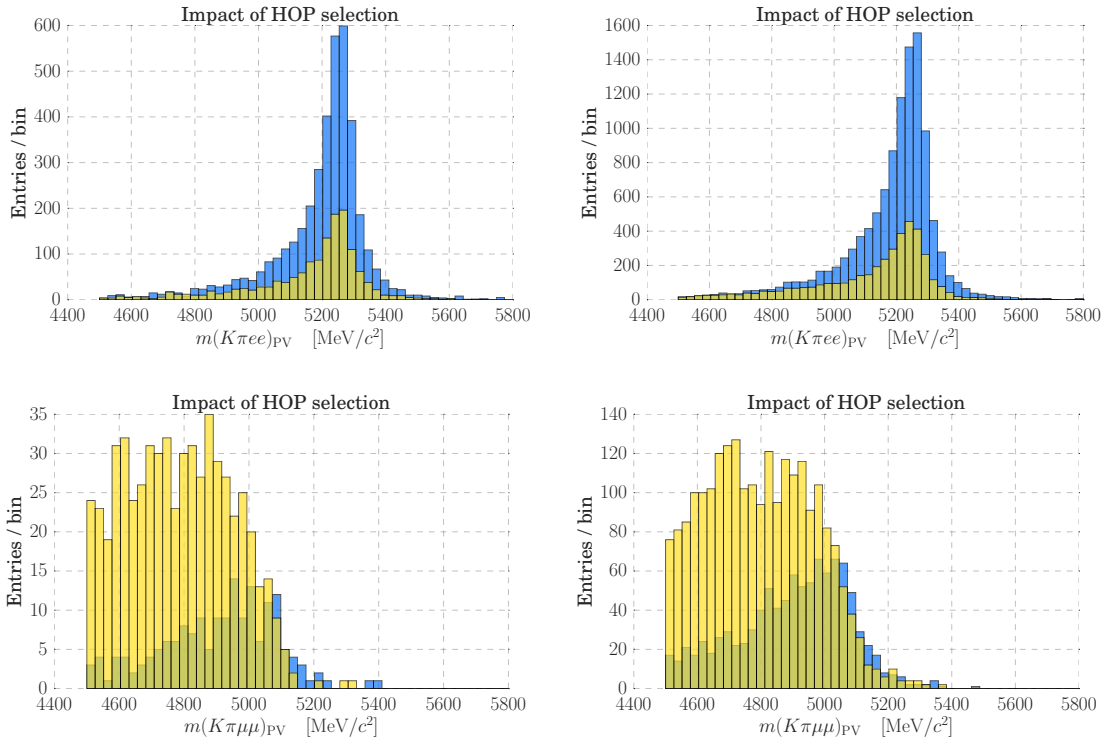


Figure 6.29: Invariant mass distribution of the $B^0 \to K^{*0}e^+e^-$ (top) and $B^+ \to K_1^+e^+e^-$ (bottom) simulated candidates in the low q^2 bin (left) and in the central q^2 bin (right) selected (blue) and discarded (yellow) by the HOP requirement.

but similar results are obtained for the other trigger categories. Only candidates passing the selection are taken into account.

The benefit of adding the HOP requirement is evaluated by performing a mass fit of the $B^0 \to K^{*0}e^+e^-$ data candidates and by computing the signal yield with and without such requirement. The fitting procedure is described in detail in Section 7.1. The fit results are shown in Fig. 6.30 for both the baseline (top) and HOP (middle) approaches. It is observed that, although the absolute number of background decays is lower when the HOP requirement is applied, the statistical uncertainty associated to the signal yield has a negligible improvement. This is due to the fact that the HOP requirement is more efficient in removing background in the low mass region than in the region where the $B^0 \to K^{*0}e^+e^-$ signal is, as shown in Fig. 6.29. For this reason, it is preferred to not include the HOP requirement in the selection.

Multivariate classifier approach

Another approach that is investigated in order to reduce the contamination from partially reconstructed decays is based on a dedicated multivariate classifier. The latter is trained to distinguish between signal decays and partially reconstructed decays. The $B^0 \to K^{*0}J/\psi$ ($\to e^+e^-$) simulation sample is used as signal, while the $B^+ \to K_1^+J/\psi$ ($\to e^+e^-$) simulation sample is used as background. The choice of using the control mode instead of the rare mode in the training is driven by the fact that these samples have a larger statistics and are expected to show the same discrepancies between

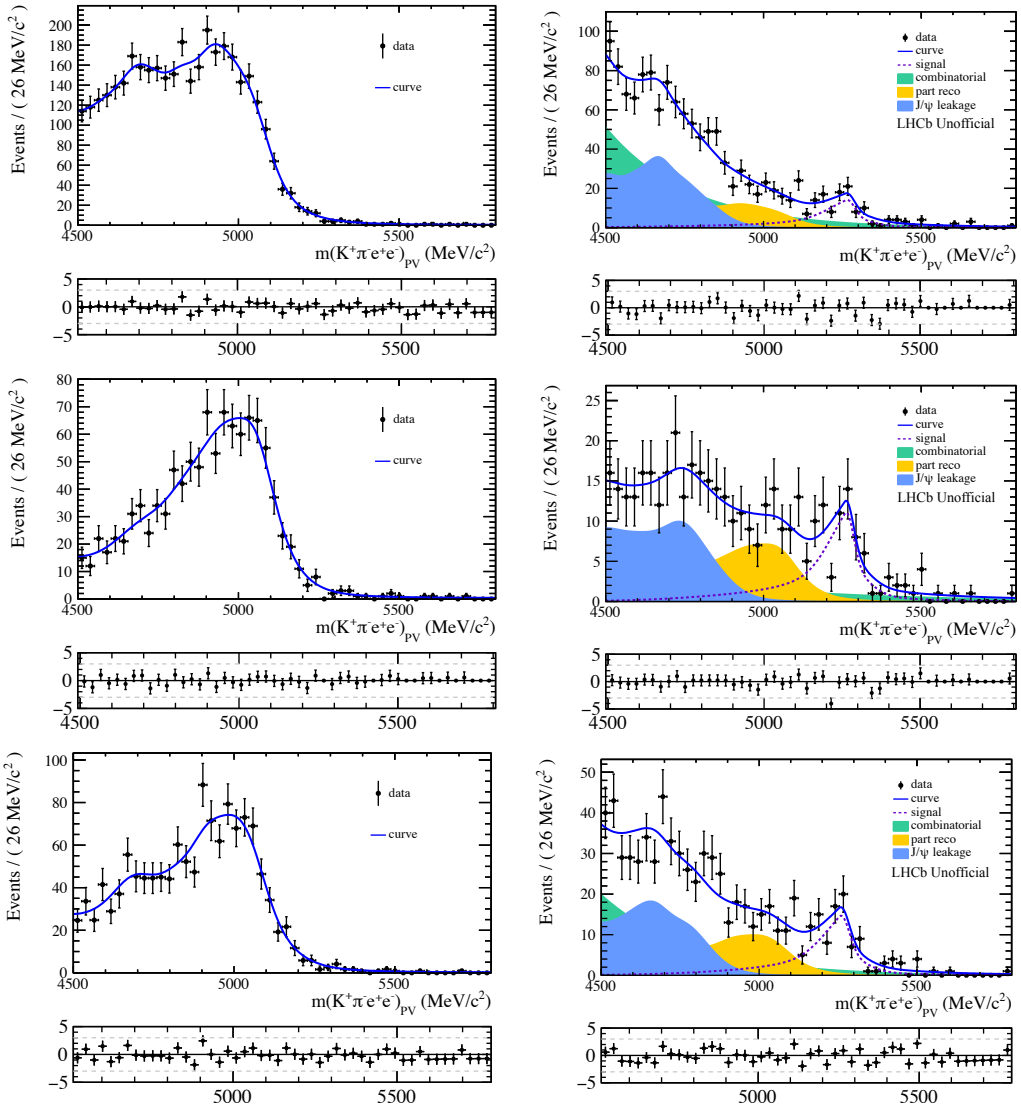


Figure 6.30: Invariant mass distribution of the $B^+ \to K_1^+ e^+ e^-$ simulated candidates (left) and $B^0 \to K^{*0} e^+ e^-$ data candidates (right) in the central q^2 bin and in the `L0E` trigger category obtained when applying the baseline selection (top) and, in addition, the HOP approach (middle) or the multivariate classifier approach (bottom).

each other as the corresponding $B^0 \to K^{*0} e^+ e^-$ and $B^+ \to K_1^+ e^+ e^-$ rare modes. In addition, the resonant modes have the advantage of allowing to determine the signal efficiency and rejection power of the trained classifier directly from the data samples.

The separation between signal and background is achieved by looking at the isolation properties of the B^0 candidates. In order to improve the rejection power of the classifier, the $\alpha_{\rm HOP}$ variable defined in the previous section is also used as discriminating variable during the training. The $m_{\rm corr}$ variable is not used since it is largely correlated to the B^0 invariant mass. In addition to these variables, the following quantities are also considered:

- $p_{\rm T}$, impact parameter χ^2, flight distance χ^2, and χ^2 of the kinematic fit of the B^0 candidate;
- impact parameter χ^2 of the leptons in the final state.

As for the HOP approach, the benefit of this strategy is assessed by performing a mass fit of the $B^0 \to K^{*0}e^+e^-$ data candidates with and without a requirement on the classifier response. The fit results are shown at the bottom and at the top of Fig. 6.30. Since the classifier removes background candidates mainly from the low mass region and given that the statistical uncertainty on the signal yield does not improve when removing these candidates, no selection on the classifier response is applied in the mass fit described in Section 7.1.

6.7.7 Multiple candidates

If multiple B^0 candidates are reconstructed in the same pp collision even after the full selection is applied, only one B^0 candidate, chosen randomly, is retained, while the others are discarded. In data, this happens for up to 2% of the $B^0 \to K^{*0}e^+e^-$ candidates and for up to 0.4% of the $B^0 \to K^{*0}\mu^+\mu^-$ candidates, depending on the q^2 bin. The corresponding values for the resonant modes are $0.3 - 0.5\%$ for $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ and 0.1% for $B^0 \to K^{*0}J/\psi$ $(\to \mu^+\mu^-)$.

6.8 Efficiency calculation

6.8.1 Absolute efficiency

The absolute efficiency, that is, the efficiency that appears in Eq. 6.1, for a given decay mode in either the S_i^+ or S_i^- category is obtained from simulation as follows:

$$\epsilon^{\pm} = \epsilon_{\text{geom}} \cdot \frac{\mathcal{N}_{\text{rec}}^{w,\pm}}{\mathcal{N}_{\text{gen}}^{q^2,\pm}}.$$

This formula corresponds to the ratio between selected and generated decays, where few considerations are in order.

The factor ϵ_{geom} is the so-called geometric efficiency of the decay mode and is due to the fact that the final-state particles of the simulated candidates are generated inside the LHCb detector acceptance only, that is, with a polar angle between 10 and 400 mrad. This is done in order to save disk space and computing time. The geometric efficiency for the relevant decay modes, which is calculated using dedicated generator-level samples, is reported in Table 6.4. The displayed values correspond to the weighted average of the geometric efficiencies for 2011 and 2012 and for the two polarities of the magnetic field.

The term $\mathcal{N}_{\text{gen}}^{q^2,\pm}$ is the total number of simulated candidates that is generated in the S_i^+ or S_i^- category and that has a true q^2 in the desired range. The true q^2 is obtained from the momenta of the B^0 and the K^{*0} candidates according to

$$q^2{}_{\text{true}} = (p^{\mu}_{B^0} - p^{\mu}_{K^{*0}})^2.$$

This is needed in order to compare the measured values of the S_i and ΔS_i observables with the predictions before the emission of final state radiation.

The term $\mathcal{N}_{\text{rec}}^{w,\pm}$ corresponds to how many generated candidates satisfy the reconstruction, trigger, and selection requirements of the analysis and have reconstructed values of the decay angles so that they belong to the S_i^+ or S_i^- category. This allows to take q^2 bin migration effects into account. The selection requirements include also a requirement on the PV-constrained B^0 invariant mass,

Table 6.4: Geometric efficiency for the decay modes used in the analysis, obtained by averaging the efficiencies of different years and polarities of the magnetic field.

Decay mode	Geometric efficiency (%)
$B^0 \to K^{*0} e^+ e^-$	16.58 ± 0.04
$B^0 \to K^{*0} \mu^+ \mu^-$	16.282 ± 0.037
$B^0 \to J/\psi\, K^{*0}$, $J/\psi \to e^+ e^-$	15.924 ± 0.038
$B^0 \to J/\psi\, K^{*0}$, $J/\psi \to \mu^+ \mu^-$	16.001 ± 0.023
$\Lambda_b^0 \to p K^- J/\psi$, $J/\psi \to e^+ e^-$	16.656 ± 0.028
$\Lambda_b^0 \to p \pi^- J/\psi$, $J/\psi \to e^+ e^-$	15.912 ± 0.027
$B^0 \to \psi(2S) K^{*0}$, $\psi(2S) \to e^+ e^-$	16.28 ± 0.04

Table 6.5: Efficiencies for the $B^0 \to K^{*0} e^+ e^-$ decay in the low and central q^2 bins, separately for the S_i^+ and S_i^- categories.

S_i	ϵ^+ low q^2 (%)	ϵ^- low q^2 (%)	ϵ^+ central q^2 (%)	ϵ^- central q^2 (%)
S_4	0.1911 ± 0.0005	0.1878 ± 0.0005	0.2058 ± 0.0005	0.2212 ± 0.0006
S_5	0.2090 ± 0.0006	0.1641 ± 0.0004	0.1961 ± 0.0005	0.2293 ± 0.0006
S_7	0.1877 ± 0.0005	0.1916 ± 0.0005	0.2158 ± 0.0006	0.2126 ± 0.0006
S_8	0.1890 ± 0.0005	0.1903 ± 0.0005	0.2132 ± 0.0006	0.2151 ± 0.0006
$A_{\rm FB}$	0.1913 ± 0.0005	0.1881 ± 0.0005	0.2209 ± 0.0006	0.2073 ± 0.0006

which has to be in the range where the fit is performed. This corresponds to the mass window from 4500 MeV/c^2 to 5800 MeV/c^2 for the rare mode $B^0 \to K^{*0} e^+ e^-$ and from 4800 MeV/c^2 to 5800 MeV/c^2 for the resonant mode $B^0 \to K^{*0} J/\psi$ ($\to e^+ e^-$). The corresponding ranges for muons are from 5150 MeV/c^2 to 5800 MeV/c^2 and from 5150 MeV/c^2 to 5400 MeV/c^2, respectively. In order to take the corrections to the simulation into account (see Section 6.6), the sum of the weights is used instead of the yield when computing $\mathcal{N}_{\rm rec}^{w,\pm}$.

For a given S_i category, the efficiency uncertainty is computed as

$$\sigma_\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)}{\mathcal{N}_{\rm gen}^{q^2}}},$$

if no weights are used, and as

$$\sigma_\epsilon = \sqrt{\frac{\epsilon(1-\epsilon)}{\mathcal{N}_{\rm gen}^{q^2}} \cdot \frac{\sum_i w_i^2}{(\sum_i w_i)^2}},$$

if the weights are taken into account.

Since the simulation sample does not reflect the fact that the statistics collected in 2012 is nearly twice that of 2011, the efficiency is evaluated separately for 2011 and 2012 and the weighted average with the abundance observed in data is used as total efficiency. The computed efficiencies for the $B^0 \to K^{*0} e^+ e^-$, $B^0 \to K^{*0} \mu^+ \mu^-$, $B^0 \to K^{*0} J/\psi$ ($\to e^+ e^-$), and $B^0 \to K^{*0} J/\psi$ ($\to \mu^+ \mu^-$) decays are reported in Tables 6.5, 6.6, 6.7, and 6.8, respectively. The values are given separately for the S_i^+ and S_i^- categories.

Table 6.6: Efficiencies for the $B^0 \to K^{*0}\mu^+\mu^-$ decay in the low and central q^2 bins, separately for the S_i^+ and S_i^- categories.

S_i	ϵ^+ low q^2 (%)	ϵ^- low q^2 (%)	ϵ^+ central q^2 (%)	ϵ^- central q^2 (%)
S_4	0.6201 ± 0.0016	0.5780 ± 0.0015	0.6611 ± 0.0016	0.6814 ± 0.0016
S_5	0.6417 ± 0.0016	0.5484 ± 0.0015	0.6374 ± 0.0015	0.7013 ± 0.0016
S_7	0.5739 ± 0.0015	0.6269 ± 0.0016	0.6904 ± 0.0016	0.6537 ± 0.0015
S_8	0.6083 ± 0.0016	0.5923 ± 0.0015	0.6734 ± 0.0016	0.6708 ± 0.0016
$A_{\rm FB}$	0.6195 ± 0.0016	0.5834 ± 0.0015	0.6892 ± 0.0016	0.6547 ± 0.0015

Table 6.7: Efficiencies for the $B^0 \to K^{*0}J/\psi$ ($\to e^+e^-$) decay, separately for the S_i^+ and S_i^- categories.

S_i	ϵ^+ (%)	ϵ^- (%)
S_4	0.2031 ± 0.0005	0.2150 ± 0.0005
S_5	0.2103 ± 0.0005	0.2095 ± 0.0005
S_7	0.2087 ± 0.0005	0.2111 ± 0.0005
S_8	0.2099 ± 0.0005	0.2099 ± 0.0005
$A_{\rm FB}$	0.2101 ± 0.0005	0.2097 ± 0.0005

In the cross-checks discussed in Section 7.2, the absolute efficiency is computed without distinguishing between the S_i^+ and S_i^- categories. This is done by using the same formulas described above, but without splitting the samples in the two categories. If the measurement is quoted in a given bin of q^2, the generation-level and reconstruction-level samples used to compute $\mathcal{N}_{\rm gen}^{q^2,\pm}$ and $\mathcal{N}_{\rm rec}^{w,\pm}$ are required to have true q^2, the first, and reconstructed q^2, the second, in the quoted interval. This is the case for the R_{K^*} measurement, which is quoted in either the low or the central q^2 bin. The situation is different for the $R_{J/\psi}$ measurement, which is expected to be unity in the full q^2 range.

6.8.2 Efficiency parametrisation

The reconstruction, trigger, and selection criteria described above modify the distribution of the signal candidates in q^2, $\cos\theta_\ell$, $\cos\theta_K$, and ϕ. This efficiency has to be taken into account in the measurement of the S_i observables, since the latter are computed as yield asymmetries between two categories of signal candidates, S_i^+ and S_i^-, whose separation is based on the reconstructed values of the decay angles. In the measurement of the S_i observables in data, this effect is taken into account as reported in Eq. (6.1). An alternative approach is used in the method validation discussed in Section 7.3.2 and is briefly described here.

Instead of assigning an efficiency to the S_i^+ and S_i^- samples, an event-by-event correction is applied. This correction is obtained from the parametrisation of the efficiency as a function of q^2, $\cos\theta_\ell$, $\cos\theta_K$, and ϕ, by comparing the distributions of the generated candidates to the resulting distributions after the reconstruction, trigger, and selection requirements are applied. The correction is tailored to candidates in different decay modes, q^2 bins, and trigger categories.

The four-dimensional distributions of the generated and selected candidates are parametrised

Table 6.8: Efficiencies for the $B^0 \to K^{*0} J/\psi\ (\to \mu^+\mu^-)$ decay, separately for the S_i^+ and S_i^- categories.

S_i	ϵ^+ (%)	ϵ^- (%)
S_4	0.8569 ± 0.0013	0.9180 ± 0.0014
S_5	0.8925 ± 0.0014	0.8908 ± 0.0014
S_7	0.8902 ± 0.0014	0.8930 ± 0.0014
S_8	0.8905 ± 0.0014	0.8928 ± 0.0014
$A_{\rm FB}$	0.8908 ± 0.0014	0.8925 ± 0.0014

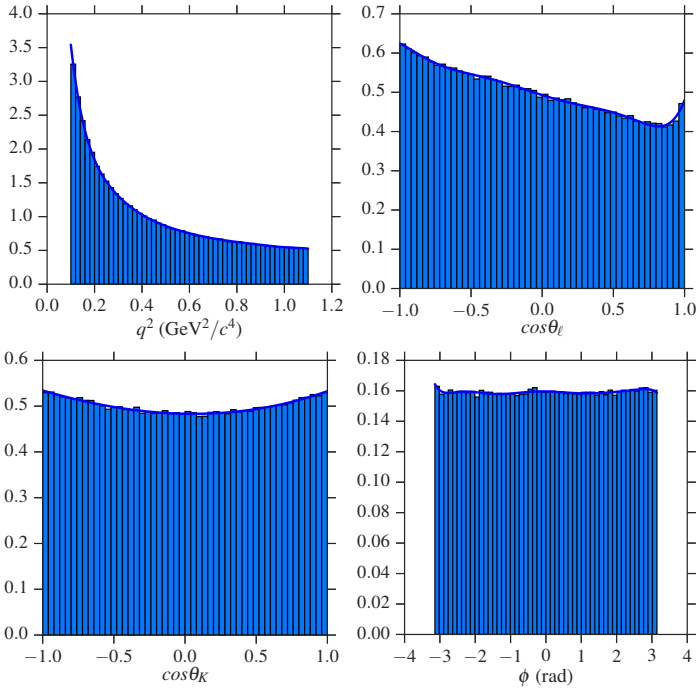


Figure 6.31: Distribution of the q^2 (top left), $\cos\theta_\ell$ (top right), $\cos\theta_K$ (bottom left), and ϕ (bottom right) variables for generation-level $B^0 \to K^{*0} e^+ e^-$ candidates in the low q^2 bin. The corresponding parametrisation in terms of Legendre polynomials is shown by the blue solid line.

according to

$$\epsilon(q^2, \cos\theta_\ell, \cos\theta_K, \phi) = \sum_{hijk} c_{hijk} L_h(q^{2\prime}) L_i(\cos\theta_\ell) L_j(\cos\theta_K) L_k(\phi'),$$

where $L_a(x)$ is the Legendre polynomial of order a in the variable x, with x having values in the range $[-1, 1]$. In this range, the Legendre polynomials are orthogonal and satisfy the following relation:

$$\int_{-1}^{1} L_n(x) L_m(x) dx = \frac{2}{2n+1} \delta_{mn},$$

where δ_{mn} is the Kronecker delta. The primed variables $q^{2\prime}$ and ϕ' in the formula above highlight the fact that, while $\cos\theta_\ell$ and $\cos\theta_K$ are already in the desired range, a change of variables is needed

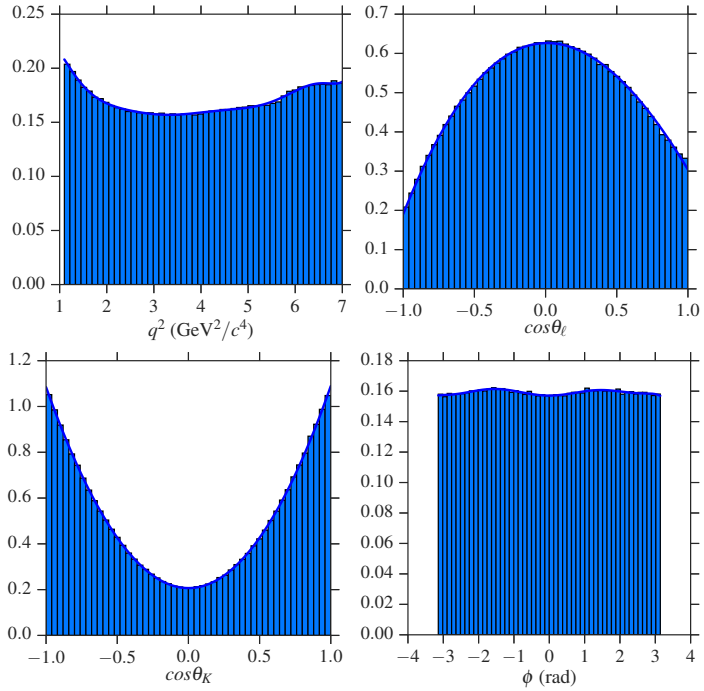


Figure 6.32: Distribution of the q^2 (top left), $\cos\theta_\ell$ (top right), $\cos\theta_K$ (bottom left), and ϕ (bottom right) variables for generation-level $B^0 \to K^{*0}e^+e^-$ candidates in the central q^2 bin. The corresponding parametrisation in terms of Legendre polynomials is shown by the blue solid line.

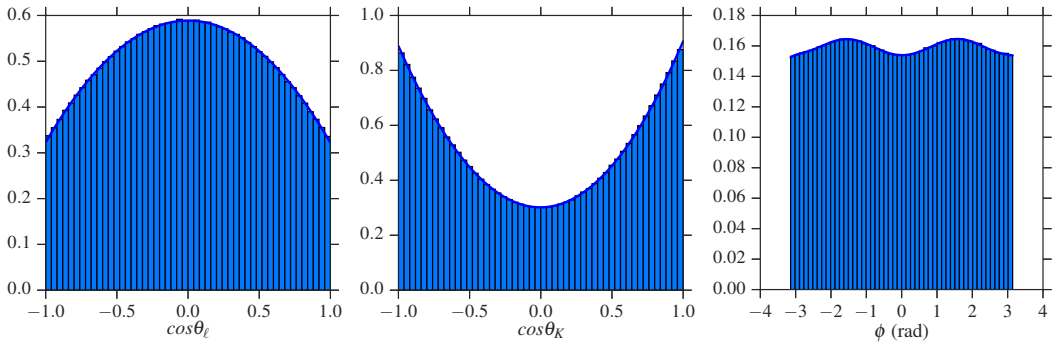


Figure 6.33: Distribution of the $\cos\theta_\ell$ (left), $\cos\theta_K$ (centre), and ϕ (right) variables for generation-level $B^0 \to K^{*0}J/\psi$ ($\to e^+e^-$) candidates. The corresponding parametrisation in terms of Legendre polynomials is shown by the blue solid line.

for q^2 and ϕ. The following change of variables is applied to q^2:

$$q^{2\prime} = \frac{2q^2 - {q^2}_{\rm min} - {q^2}_{\rm max}}{{q^2}_{\rm max} - {q^2}_{\rm min}},$$

where ${q^2}_{\rm min}$ and ${q^2}_{\rm max}$ are $0.1\,\text{GeV}^2/c^4$ and $1.1\,\text{GeV}^2/c^4$, for the low q^2 bin, and $1.1\,\text{GeV}^2/c^4$ and $7.0\ \text{GeV}^2/c^4$, for the central q^2 bin. The values of ϕ are instead transformed according to

$$\phi' = \frac{\phi}{\pi},$$

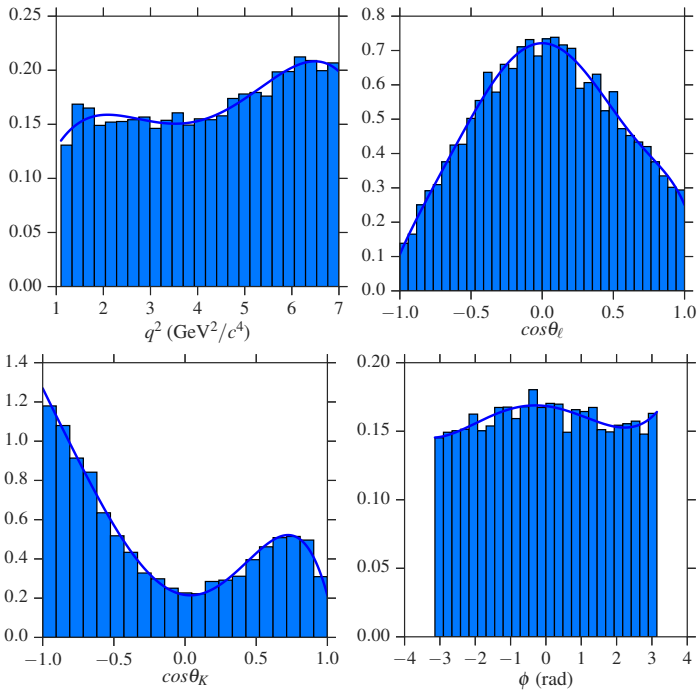


Figure 6.34: Distribution of the q^2 (top left), $\cos\theta_\ell$ (top right), $\cos\theta_K$ (bottom left), and ϕ variables for reconstruction-level $B^0 \to K^{*0}e^+e^-$ candidates in the central q^2 bin. The corresponding parametrisation in terms of Legendre polynomials is shown by the blue solid line.

since ϕ is originally defined between $-\pi$ and π.

The coefficients c_{hijk} are computed by means of a principal moment analysis. The corrections of the simulation discussed in Section 6.6.2 and 6.6.3 are taken into account when computing the coefficients. The maximum order of the Legendre polynomials in q^2, $\cos\theta_\ell$, $\cos\theta_K$, and ϕ corresponds to the lowest order that allows a good description of the distributions of the simulated candidates. The same orders are used for the $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ modes. In particular, Legendre polynomials of order up to the 10, 8, 4, and 10 (10, 4, 4, 10) are used to parametrise the q^2, $\cos\theta_\ell$, $\cos\theta_K$, and ϕ distributions of the generation-level samples in the low (central) q^2 bin, while Legendre polynomials of order up to the 5, 8, 6, and 5, are used to parametrise the distributions of the reconstruction-level samples, using the same orders for the low and central q^2 bins. The different choice for generation-level and reconstruction-level samples is due to the lower statistics in the latter, which makes the parametrisation too sensitive to statistical fluctuations if higher orders are taken into account. For simplicity, the same maximum orders are adopted in the parametrisation of the reconstruction-level distributions of all decay modes, q^2 bins, and trigger categories.

As a cross-check, the coefficients c_{hijk} of the generation-level $B^0 \to K^{*0}e^+e^-$ candidates in the central q^2 bin and the generation-level $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ candidates are recomputed by requiring $\cos\theta_\ell$ and ϕ' to have distributions symmetric around zero. This allows to set to zero the coefficients corresponding to an odd order of $\cos\theta_\ell$, ϕ', or both. The non-vanishing coefficients are then compared to those obtained without the symmetry requirement and are found to be compatible.

The efficiency is taken into account by weighting each simulated candidate by $w^e = \epsilon^e_{\rm gen}/\epsilon^e_{\rm rec}$, with $\epsilon^e_{\rm gen}$ and $\epsilon^e_{\rm sel}$ being the efficiencies of the generated and selected distributions. The sum of these acceptance weights is then used in the method validation described in Section 7.3.2. The acceptance

weights could be taken into account in the fit to the data candidates described in Section 7.1 instead of dividing the signal yields obtained from the fit by the absolute efficiency. However, it is preferred to follow the approach described in Eq. (6.1), since the acceptance weights have a large uncertainty due to the limited statistics of the reconstruction-level samples from which they are obtained.

The parametrisations of the q^2, $\cos\theta_\ell$, $\cos\theta_K$, and ϕ distributions of the generation-level $B^0 \to K^{*0}e^+e^-$ candidates are shown in Figs. 6.31 and 6.32, respectively for the low and central q^2 bins. The pararametrisation of the generation-level $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ candidates is shown in Fig. 6.33. The distributions after the reconstruction, trigger, and selection criteria are applied are shown in Fig. 6.34 for the $B^0 \to K^{*0}e^+e^-$ candidates in the `L0E` trigger category. Similar plots are availabe for the other two trigger categories.

CHAPTER 7

Mass fits and measurement of the S_i and ΔS_i observables

This chapter describes the fit to the invariant mass of the $B^0 \to K^{*0}\ \ell^+\ell^-$ candidates in simulation and data and presents the results of the S_i and ΔS_i measurements for decays to electrons and to muons in the low and central q^2 regions. The method validation and the cross-checks that are performed are also discussed. The chapter ends with a discussion of the implications of the results and an overview of the future prospects.

7.1 Mass fits

The measurement of the S_i and ΔS_i observables requires to determine the signal yields and the efficiencies of Eq. (6.1). The strategy to compute the efficiencies is described in Section 6.8, while the procedure to determine the signal yields is discussed here.

The signal yields are obtained from an unbinned extended maximum likelihood fit to the invariant mass of the $B^0 \to K^{*0}\ \ell^+\ell^-$ candidates in data. In particular, the PV-constrained B^0 invariant mass is used in the fit. A fit of the full data sample, without distinguishing between the S_i^+ and S_i^- categories, is performed in order to validate the choice of the model. The mass fit is performed independently for the low and central q^2 bins. Some of the parameters of the model are obtained from simulation and, for the decay to electrons, also from a fit of the PV-constrained B^0 invariant mass distribution of the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ data candidates. Since some parameters of the model are expected to be common between the rare and the resonant mode, this approach allows to use the large statistics available in the resonant mode to compute the best-fit values of the rare mode. For the decay to electrons, the three trigger categories `L0E`, `L0H`, and `L0TIS` are combined together in the fit to the data.

The corrections of the simulation are included in the mass fit by weighting each simulated candidate by the total weight resulting from the corrections described in Sections 6.6.2 and 6.6.3.

The weights are included in the likelihood as follows:

$$\mathcal{L} = -\sum_{e=1}^{N} w_e \times \log \mathcal{P}(\vec{\Omega}_e, m_e),$$

where the sum runs over the number of signal candidates in the sample. The parameter uncertainties are estimated using an approximate method that guarantees the correct coverage in the case of weighted fits. The method consists in computing a corrected covariance matrix V' as follows:

$$V' = VC^{-1}V,$$

where V and C are the covariance matrices obtained from two different weighted fits, one using the weights w_e and the other using the squared weights w_e^2.

7.1.1 Electron mode

The signal, which corresponds to $B^0 \to K^{*0}e^+e^-$ decays, is described by the sum of a DCB and a Gaussian distribution:

$$\text{PDF}_{\text{sig}} = f_{\text{DCB}} \cdot \text{DCB}(\mu_{\text{DCB}}, \sigma_{\text{DCB}}, n_1, n_2, \alpha_1, \alpha_2, f) + (1 - f_{\text{DCB}})\text{G}(\mu_{\text{G}}, \sigma_{\text{G}}),$$

where f_{DCB} is the fraction of the DCB contribution with respect to the total PDF, μ_{DCB}, σ_{DCB}, n_1, n_2, α_1, α_2, and f are the parameters of the DCB, as described in Section 6.6.2, and μ_{G} and σ_{G} are the mean and width of the Gaussian distribution. The same model is used to parametrise the mass distribution of the $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ candidates. The DCB allows to parametrise the radiative tail of the decays to electrons and the effect of the bremsstrahlung overcorrection, which is visible in the high-mass sideband.

The results of the fit to the $B^0 \to K^{*0}e^+e^-$ simulated candidates in the low and central q^2 bins are shown in Fig. 7.1, while the results obtained from the $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ simulated candidates are shown in Fig. 7.2. In both cases, the fit is performed separately for the three trigger categories.

When fitting the data, the tail parameters of the DCB (that is, n_1, n_2, α_1, and α_2), the mean and width of the DCB and of the Gaussian distribution (that is, μ_{DCB}, σ_{DCB}, μ_{G}, and σ_{G}), and the relative fraction between the two CBs (that is, f) are fixed to the corresponding values obtained from simulation, independently for each trigger category and separately for the low and central q^2 bins of the rare mode and the single q^2 bin of the resonant mode. In order to be less sensitive to detector effects that are not well reproduced in the simulation, the values of the mean and width parameters are allowed to shift and scale with respect to the values obtained from simulation, respectively. The shift of the mean and the scale factor of the width, which are common to the DCB and the Gaussian distribution, are obtained from the data fit of the resonant mode, since the statistics of the rare mode does not allow to extract this information reliably. Finally, the relative fraction between the DCB and the Gaussian distribution, f_{DCB}, is also obtained from the data fit of the resonant mode.

In addition to the signal, the following contributions are modelled in the invariant mass of the rare mode:

- combinatorial background, described by an exponential distribution;

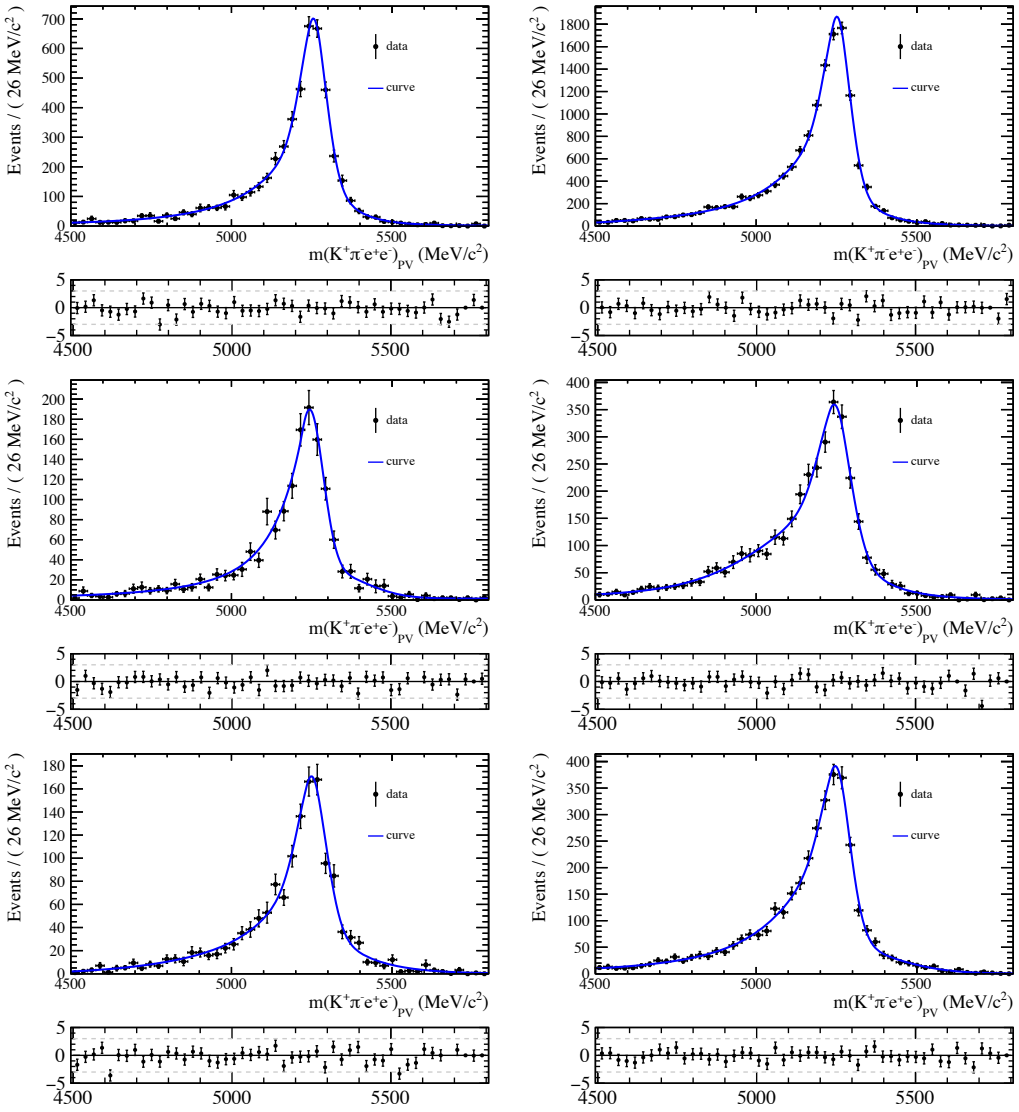


Figure 7.1: Distribution of the $B^0 \to K^{*0}e^+e^-$ simulated candidates (black points with error bars) and corresponding weighted fit (solid blue line) for the low (left) and central (right) q^2 bins and for the `L0E` (top), `L0H` (middle), and `L0TIS` (bottom) trigger categories. The residuals are also shown.

- background due to partially reconstructed decays, described by a non-parametric PDF based on KDE;
- for the central q^2 bin only, background due to $B^0 \to K^{*0}J/\psi\ (e^+e^-)$ decays leaking in the central q^2 region, described by a non-parametric PDF similar to the one above.

The template describing the background due to partially reconstructed decays is obtained from simulation. This is done by using a sample consisting of a mixture of $B^+ \to K_1^+e^+e^-$ and $B^+ \to K_2^*(1430)^+e^+e^-$ decays. A weight is assigned to the candidates of each decay mode in order to reproduce the proportion of $K_1^+ : K_2^{*+}$ corresponding to $2:1$, which is based on previous studies [34]. A weighted unbinned extended maximum likelihood fit is then performed using the non-parametric PDF. The fit results are shown in Fig. 7.3, separately for the low and central q^2 bins and for the

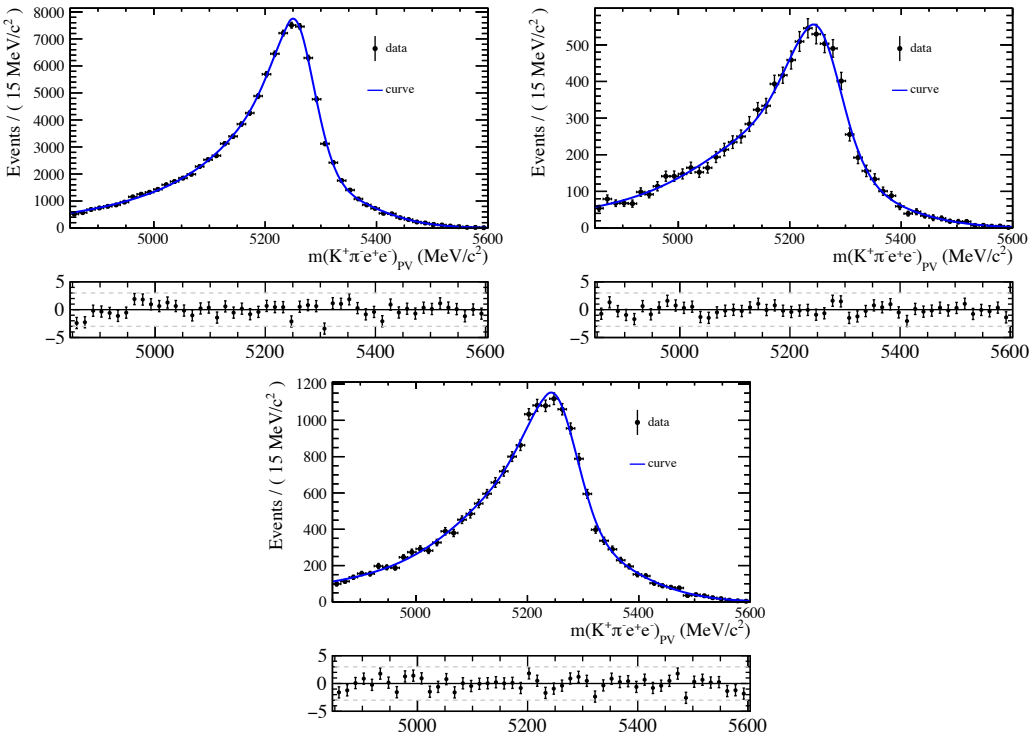


Figure 7.2: Distribution of the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ simulated candidates (black points with error bars) and corresponding weighted fit (solid blue line) for the `L0E` (top left), `L0H` (top right), and `L0TIS` (bottom) trigger categories. The residuals are also shown.

three trigger categories. A second method is planned to be investigated during the evaluation of the systematic uncertainty. In this second method, the $B^+ \to K_1^+ e^+ e^-$ simulation sample with a flat $(K\pi\pi)$ invariant mass is reweighted according to the $K^+\pi^+\pi^-$ invariant mass distribution observed in data and the resulting distribution is fitted with the non-parametric PDF.

The template describing the background due to $B^0 \to K^{*0} J/\psi\ (e^+e^-)$ decays leaking in the central q^2 region is obtained in a similar way by fitting the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ simulation sample with a non-parametric PDF. The fit results are shown in Fig. 7.4, separately for the three trigger categories.

When fitting the data, the yield of this background contribution, $\mathcal{N}_{J/\psi\ \text{leak}}$, is constrained to the yield of the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ signal, $\mathcal{N}_{J/\psi\ \text{sig}}$, which is left free to vary in the fit of the resonant mode. The constraint consists in requiring that

$$\mathcal{N}_{J/\psi\ \text{leak}} = \mathcal{N}_{J/\psi\ \text{sig}} \cdot \frac{\mathcal{N}^{\text{MC}}_{J/\psi \text{leak}}}{\mathcal{N}^{MC}_{J/\psi\ \text{sig}}},$$

where $\mathcal{N}^{\text{MC}}_{J/\psi \text{leak}}$ and $\mathcal{N}^{MC}_{J/\psi\ \text{sig}}$ are the yields of $B^0 \to K^{*0} J/\psi\ (e^+e^-)$ simulated decays leaking in the central q^2 region and falling in the q^2 region corresponding to the resonant mode, respectively.

As a result, the model for the $B^0 \to K^{*0} e^+ e^-$ candidates in the low q^2 bin can be written as

$$\text{PDF}_{\text{low}} = \mathcal{N}_{\text{sig}} \cdot \text{PDF}_{\text{sig}} + \mathcal{N}_{\text{comb}} \cdot \text{PDF}_{\text{comb}} + \mathcal{N}_{\text{part reco}} \cdot \text{PDF}_{\text{part reco}},$$

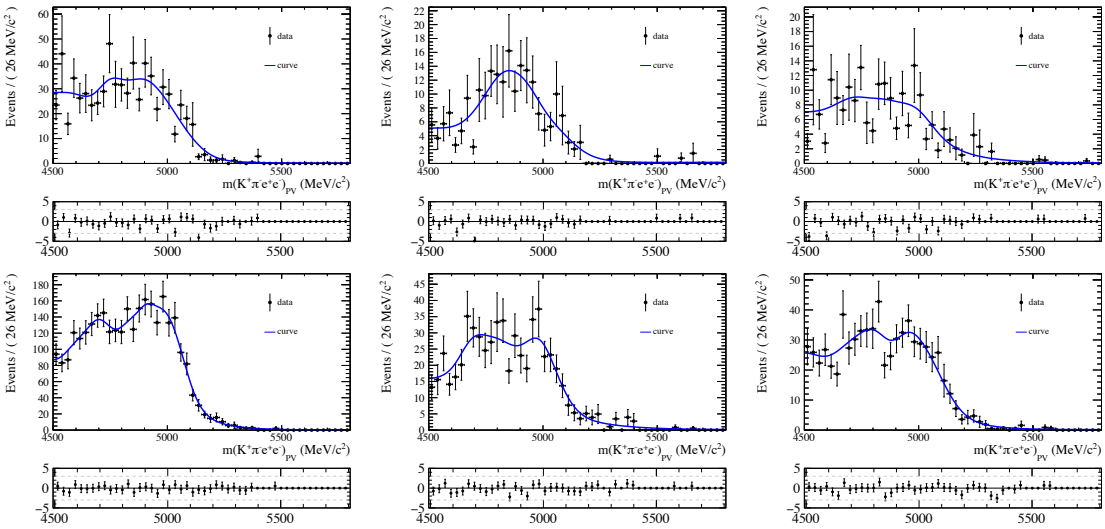


Figure 7.3: Distribution of the background due to partially reconstructed decays in the low (top) and central (bottom) q^2 bin (black points with error bars) and corresponding weighted fit (solid blue line) for the `L0E` (left), `L0H` (center), and `L0TIS` (right) trigger categories. The residuals are also shown.

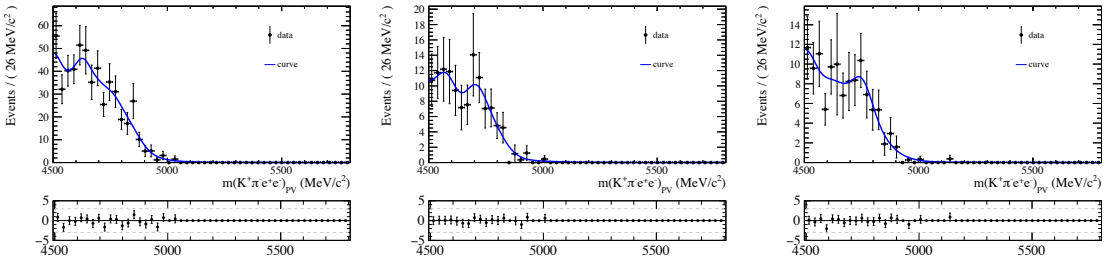


Figure 7.4: Distribution of the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ simulated candidates leaking in the central q^2 bin (black points with error bars) and corresponding weighted fit (solid blue line) for the `L0E` (left), `L0H` (center), and `L0TIS` (right) trigger categories. The residuals are also shown.

while the one for the $B^0 \to K^{*0} e^+e^-$ candidates in the central q^2 bin can be written as

$$\mathrm{PDF}_{\mathrm{central}} = \mathcal{N}_{\mathrm{sig}} \cdot \mathrm{PDF}_{\mathrm{sig}} + \mathcal{N}_{\mathrm{comb}} \cdot \mathrm{PDF}_{\mathrm{comb}} + \mathcal{N}_{\mathrm{part\ reco}} \cdot \mathrm{PDF}_{\mathrm{part\ reco}} + \mathcal{N}_{\mathrm{J/\psi\ leak}} \cdot \mathrm{PDF}_{\mathrm{J/\psi\ leak}}.$$

Due to the low statistics available in the data sample of the rare mode, all components of the formulas above except the combinatorial background consist of the weighted sum of the templates of the three trigger categories. The relative abundance of each trigger category is obtained from simulation. The impact of this choice on the results is taken into account as systematic uncertainty.

In the fit to the control mode, the following contributions are considered in addition to the signal:

- $B^0_s \to K^{*0} J/\psi\ (\to e^+e^-)$ decays, described by the same model of the signal, with a mass shift corresponding to the difference between the B^0_s and B^0 nominal masses, as reported in Ref. [15];
- combinatorial background, described by an exponential distribution.

Three additional background contributions are investigated but not included in the model since their yields are found to be negligible with respect to the total yield. These contributions are

Table 7.1: Results of the fit to the $B^0 \to K^{*0}e^+e^-$ data candidates in the low and central q^2 bins.

Component	Yield at low q^2	Yield at central q^2
$B^0 \to K^{*0}e^+e^-$	78^{+16}_{-15}	188^{+27}_{-27}
combinatorial background	204^{+40}_{-50}	1284^{+90}_{-80}
partially reconstructed decays	95^{+39}_{-36}	276^{+70}_{-70}

Table 7.2: Results of the fit to the $B^0 \to K^{*0}J/\psi\ (\to e^+e^-)$ data candidates.

Component	Yield in `L0E`	Yield in `L0H`	Yield in `L0TIS`
$B^0 \to K^{*0}J/\psi\ (\to e^+e^-)$	34901^{+190}_{-190}	2956^{+70}_{-80}	9115^{+100}_{-90}
combinatorial background	14^{+28}_{-9}	19^{+60}_{-60}	$1.1^{+1.1}_{-1.1}$

the background due to partially reconstructed decays, which is vetoed by requiring the PV- and J/ψ-constrained B^0 invariant mass of the candidates to be larger than 5150 MeV/c^2, the background due to Λ_b^0 decays, which is vetoed by the selection requirements in Section 6.7.2, and the background due to $B^0 \to K^{*0}\psi(2S)\ (\to e^+e^-)$ decays leaking in the q^2 region between 7.0 and 11.0 GeV$^2/c^4$. Due to the much larger statistics, the combination of the three trigger categories is not needed when fitting the data sample of the resonant mode.

In summary, the following parameters are left free to vary in the fit to the $B^0 \to K^{*0}J/\psi\ (\to e^+e^-)$ data candidates:

- f_{DCB}, the shift of the mean, and the scale factor of the width (three parameters);
- the signal yield of the resonant mode (one parameter);
- the yield and slope of the combinatorial background (two parameters).

The following parameters are left free to vary in the fit to the $B^0 \to K^{*0}e^+e^-$ data candidates, independently for the low and central q^2 bins:

- the signal yields (two parameters);
- the yield and slope of the combinatorial background (four parameters);
- the yield of the background due to partially reconstructed decays (two parameters).

The results of the fit to the data are shown in Fig. 7.5 for the rare mode and in Fig. 7.6 for the resonant mode. The signal and background yields obtained from the fit are shown in Table 7.1 for the rare mode and in Table 7.2 for the resonant mode.

7.1.2 Muon mode

The $B^0 \to K^{*0}\mu^+\mu^-$ data candidates are fitted with the sum of a DCB for the signal and an exponential distribution for the combinatorial background. Due to the good mass resolution of the muon mode, the contamination from $B^0 \to K^{*0}J/\psi\ (\to \mu^+\mu^-)$ decays leaking to the central q^2 bin is

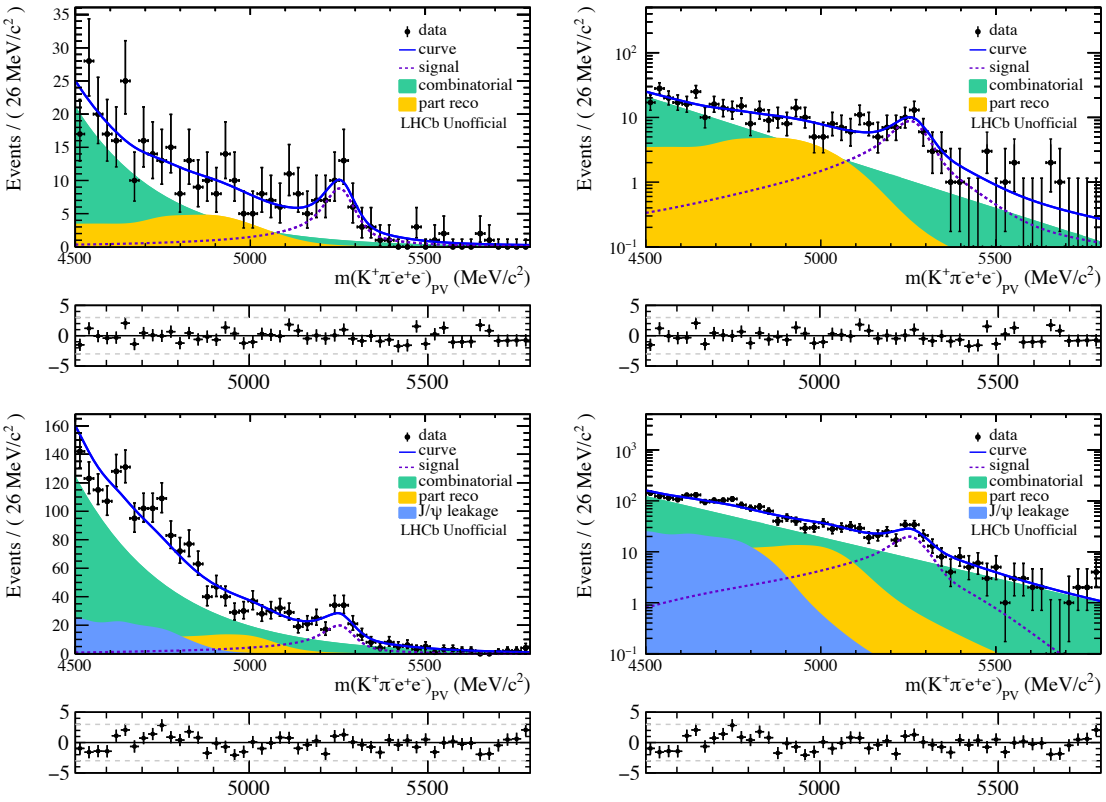


Figure 7.5: Distribution of the $B^0 \to K^{*0}e^+e^-$ data candidates (black points with error bars) and corresponding fit of the combined `L0E`, `L0H`, and `L0TIS` trigger categories (solid blue line) for the low (top) and central (bottom) q^2 bins, in normal scale (left) and log scale (right). The residuals are also shown.

negligible and is hence not modelled in the fit. The background due to partially reconstructed decays is not taken into account in the model since it is expected to be negligible in the mass window from 5150 to 5800 MeV/c^2 where the fit is performed.

The template for fitting the data is partially obtained from simulation, similarly to what discussed in Section 7.1.1. The parameters fixed from the fit to the $B^0 \to K^{*0}\mu^+\mu^-$ simulated candidates are those corresponding to the tails of the DCB, namely α_1, α_2, n_1, n_2, and the relative fraction of one CB with respect to the other, that is, f. When fitting the data sample, the mean and width of the DCB are allowed to respectively shift and scale with respect to the values obtained from simulation. An unbinned extended maximum likelihood fit is performed to determine the signal and background yields. The fit results are shown in Fig. 7.7 for simulation and in Fig. 7.8 for data.

The resonant mode is fitted with the same template used to parametrise the mass distribution of the $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ candidates. The fit results for the $B^0 \to K^{*0}J/\psi$ $(\to \mu^+\mu^-)$ simulated and data candidates are shown in Fig. 7.9.

The signal and background yields obtained from the fit are shown in Table 7.3 for the rare mode and in Table 7.4 for the resonant mode.

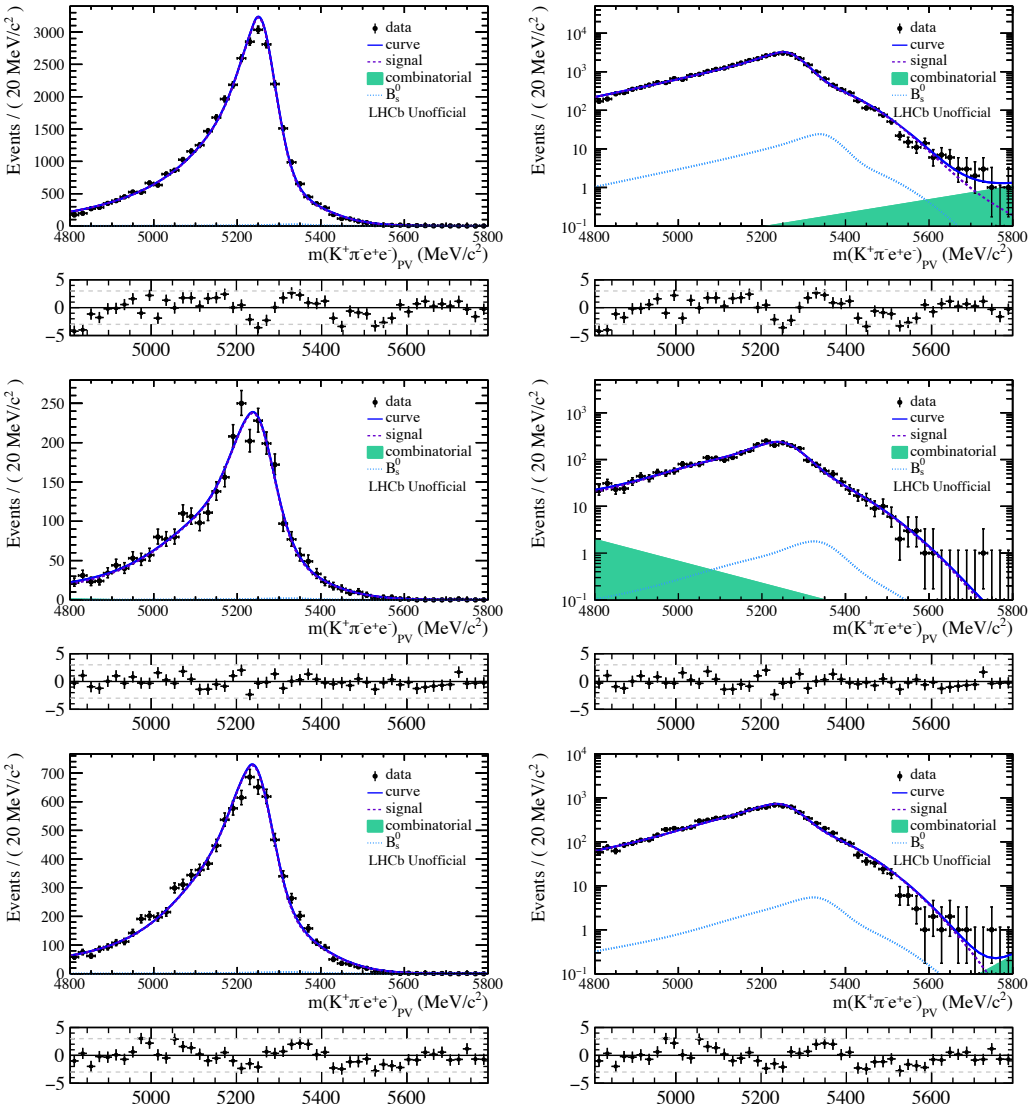


Figure 7.6: Distribution of the $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ data candidates (black points with error bars) and corresponding fit (solid blue line) for the `L0E` (top), `L0H` (middle), and `L0TIS` (right) trigger categories with normal (left) and log (right) scale. The residuals are also shown.

7.2 Cross-checks

Two cross-checks are performed. The first is a measurement of $R_{J/\psi}$ in the full q^2 spectrum and the second is a measurement of R_{K^*} in the low and central q^2 bins. The results are compared to the latest measurement by the LHCb collaboration [34].

7.2.1 Measurement of $R_{J/\psi}$

The first cross-check consists in measuring the ratio between the $B^0 \to K^{*0} J/\psi\ (\to \mu^+\mu^-)$ and $B^0 \to K^{*0} J/\psi\ (\to e^+e^-)$ branching fractions in the full spectrum of q^2. This is expected to be equal

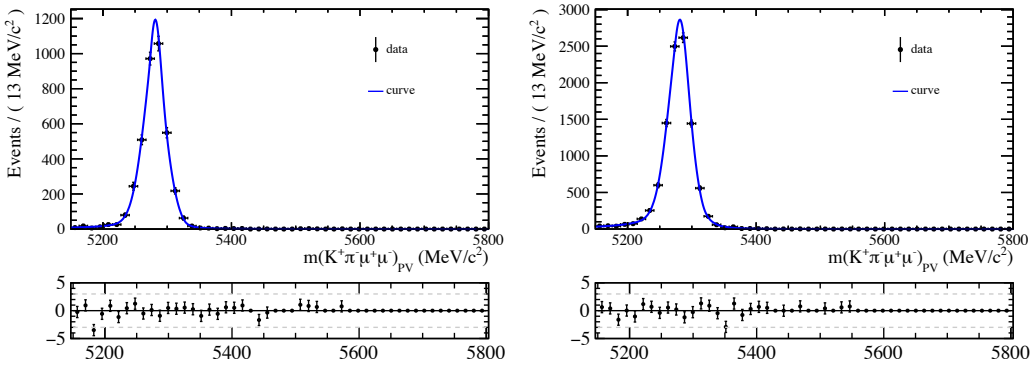


Figure 7.7: Distribution of the $B^0 \to K^{*0}\mu^+\mu^-$ simulated candidates (black points with error bars) and corresponding weighted fit (solid blue line) for the low (left) and central (right) q^2 bins. The residuals are also shown.

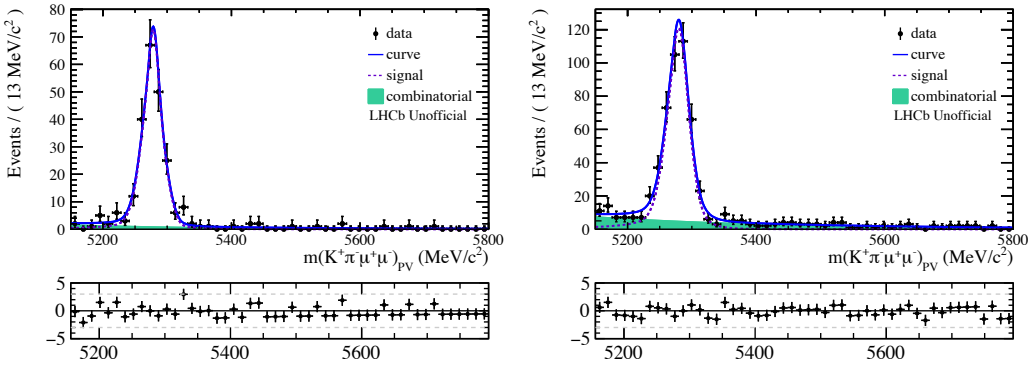


Figure 7.8: Distribution of the $B^0 \to K^{*0}\mu^+\mu^-$ data candidates (black points with error bars) and corresponding fit (solid blue line) for the low (left) and central (right) q^2 bins. The residuals are also shown.

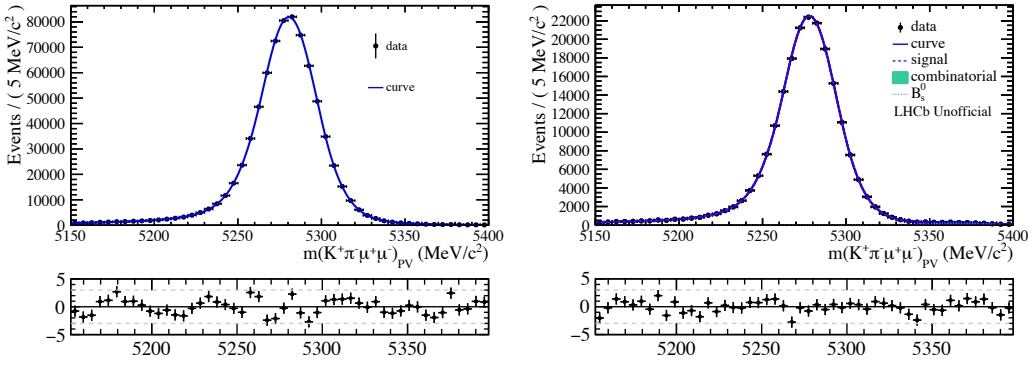


Figure 7.9: Distribution of the $B^0 \to K^{*0}J/\psi$ $(\to\mu^+\mu^-)$ candidates (black points with error bars) and corresponding fit (solid blue line) for the simulated (left) and data (right) candidates in the `LOM` trigger category. The residuals are also shown.

to unity. The value of $R_{J/\psi}$ is computed according to

$$R_{J/\psi} = \frac{\mathcal{B}(B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-))}{\mathcal{B}(B^0 \to K^{*0}J/\psi(\to e^+e^-))} = \frac{\mathcal{N}(B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-))}{\mathcal{N}(B^0 \to K^{*0}J/\psi(\to e^+e^-))} \cdot \frac{\epsilon_{J/\psi\to ee}}{\epsilon_{J/\psi\to\mu\mu}},$$

Table 7.3: Results of the fit to the $B^0 \to K^{*0}\mu^+\mu^-$ data candidates in the low and central q^2 bins.

Component	Yield at low q^2	Yield at central q^2
$B^0 \to K^{*0}\mu^+\mu^-$	211^{+16}_{-16}	420^{+24}_{-23}
combinatorial background	32^{+9}_{-8}	162^{+17}_{-16}

Table 7.4: Results of the fit to the $B^0 \to K^{*0}J/\psi\ (\to\mu^+\mu^-)$ data candidates.

Component	Yield in LOM
$B^0 \to K^{*0}J/\psi\ (\to\mu^+\mu^-)$	206693^{+1800}_{-500}
combinatorial background	466^{+220}_{-1800}

where $\mathcal{N}(B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-))$ and $\mathcal{N}(B^0 \to K^{*0}J/\psi(\to e^+e^-))$ are the yields of the $B^0 \to K^{*0}J/\psi\ (\to\mu^+\mu^-)$ and $B^0 \to K^{*0}J/\psi\ (\to e^+e^-)$ decays and $\epsilon_{J/\psi\to ee}$ and $\epsilon_{J/\psi\to\mu\mu}$ are the absolute efficiencies of the two decay modes. It is important to note that this is a stringent test, since some of the systematic uncertainties do not simplify in the ratio between muons and electrons, as they do in the second cross-check.

The yields are obtained similarly to what discussed in the previous section. The absolute efficiencies are evaluated from the relevant simulation samples as described in Section 6.8.1, without any requirement on the true q^2 of the generated candidates. The results are shown in Table 7.5 and have to be compared with $R_{J/\psi} = 1.043 \pm 0.006\,(\text{stat.}) \pm 0.045\,(\text{syst.})$ measured by LHCb [34]. One important consideration is that only the statistical uncertainty is taken into account in the $R_{J/\psi}$ cross-check. According to the LHCb measurement, this is expected to be a minor contribution compared to the much larger systematic uncertainty. Therefore, the evaluation of the systematic uncertainty is needed in order to properly compare the two results. In addition, a larger systematic uncertainty is expected for the measurement in the LOTIS trigger category. This is due to the trigger efficiency correction, that is obtained from a sample with limited statistics. Assuming the same systematic uncertainty for this cross-check and the LHCb measurement, the obtained values show that the mass fit procedure and the computation of the absolute efficiencies are sound.

7.2.2 Measurement of R_{K^*}

The value of R_{K^*} is computed according to the following double ratio:

$$R_{K^*} = \frac{\mathcal{N}(B^0 \to K^{*0}\mu^+\mu^-)}{\mathcal{N}(B^0 \to K^{*0}e^+e^-)} \cdot \frac{\epsilon_{ee}}{\epsilon_{\mu\mu}} \cdot \frac{\mathcal{N}(B^0 \to K^{*0}J/\psi(\to e^+e^-))}{\mathcal{N}(B^0 \to K^{*0}J/\psi(\to \mu^+\mu^-))} \cdot \frac{\epsilon_{J/\psi\to\mu\mu}}{\epsilon_{J/\psi\to ee}}.$$

In the formula, $\mathcal{N}(B^0 \to K^{*0}\mu^+\mu^-)$ and $\mathcal{N}(B^0 \to K^{*0}e^+e^-)$ are the yields of the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays, ϵ_{ee} and $\epsilon_{\mu\mu}$ are the absolute efficiencies of the two decay modes, and the remaining symbols are as described in Section 7.2.1. The double ratio approach helps to reduce the systematic uncertainties that are common to rare and resonant modes.

The yield of each decay mode is obtained from an unbinned extended maximum likelihood fit of the relevant data sample, with the templates described in this chapter. The absolute efficiencies are

Table 7.5: Results of the $R_{J/\psi}$ cross-check. Only the statistical uncertainty is shown.

Trigger	$R_{J/\psi}$
L0E	1.073 ± 0.007
L0H	1.095 ± 0.031
L0TIS	0.794 ± 0.009
Combined	0.972 ± 0.005

Table 7.6: Results of the R_{K^*} cross-check in the low and central q^2 bins.

Variable	Low q^2	Central q^2
$\mathcal{N}_{\text{sig}}$	82 ± 16	171 ± 24
$\mathcal{N}_{\text{sig}}^{\text{LHCb}}$ (L0E)	55^{+9}_{-8}	67^{+10}_{-10}
$\mathcal{N}_{\text{sig}}^{\text{LHCb}}$ (L0H)	13^{+5}_{-5}	19^{+6}_{-5}
$\mathcal{N}_{\text{sig}}^{\text{LHCb}}$ (L0TIS)	21^{+5}_{-4}	25^{+7}_{-6}
σ_{rel}	0.20	0.15
$\sigma_{\text{rel}}^{\text{LHCb}}$	0.17	0.16

computed from the corrected simulation samples as detailed in Section 6.8. Since the R_{K^*} value is measured in bins of q^2, a selection on the true q^2 of the generated candidates is applied in the efficiency calculation. The results are found to be compatible with those measured in Ref. [34] for both bins of q^2. It is important to note, however, that due to the different definition of the q^2 bins between this analysis and the measurement in Ref. [34], the R_{K^*} values are not expected to be exactly the same.

The relative uncertainties on the values measured in this cross-check are shown in Table 7.6, where they are compared to the corresponding uncertainties of the measurement performed at LHCb [34]. The table shows also the signal yields obtained from the fit to the $B^0 \to K^{*0} e^+ e^-$ data candidates in this cross-check and in the measurement in Ref. [34]. Since the R_{K^*} measurement performed in this thesis is a consistency check with respect to previous LHCb results, the measured central value is not shown.

The different choice of the low q^2 bin, which starts at $0.1\,\text{GeV}^2/c^4$ instead of $0.045\,\text{GeV}^2/c^4$ as in the measurement of Ref. [34], reduces the statistics available but allows to compare muons and electrons in a kinematic region in which their behaviour is expected to be universal [133]. The signal yield in the central q^2 bin is significantly larger than in the measurement of Ref. [34] due to the choice of extending the bin up to $7.0\,\text{GeV}^2/c^4$ instead of $6.0\,\text{GeV}^2/c^4$. This is possible because of the selection on the constrained q^2 discussed in Section 6.7.4.

7.3 Measurement of the S_i and ΔS_i observables

The S_i observables are measured by simultaneously fitting the candidates in the S_i^+ and S_i^- categories. The same models described in Section 7.1 are used in the fit. In particular, the data candidates in the S_i^+ and S_i^- categories are fitted simultaneously with the same model, except for few parameters

Table 7.7: Values of the S_i observables used in the generation of the $B^0 \to K^{*0}\mu^+\mu^-$ pseudo-experiments in the SM scenario, compared to the values measured with the counting method approach.

S_i	Low q^2		Central q^2	
	Generated	Measured	Generated	Measured
S_4	-0.065	-0.09 ± 0.11	0.137	0.19 ± 0.08
S_5	0.193	0.25 ± 0.09	-0.186	-0.27 ± 0.07
S_7	0.000	0.01 ± 0.09	0.000	0.01 ± 0.07
S_8	0.000	-0.01 ± 0.11	0.000	0.00 ± 0.08
$A_{\rm FB}$	0.111	0.12 ± 0.06	-0.049	-0.08 ± 0.05

Table 7.8: Values of the S_i observables used in the generation of the $B^0 \to K^{*0}\mu^+\mu^-$ pseudo-experiments in the NP scenario, compared to the values measured with the counting method approach.

S_i	Low q^2		Central q^2	
	Generated	Measured	Generated	Measured
S_4	-0.036	-0.06 ± 0.11	0.137	0.18 ± 0.08
S_5	0.242	0.29 ± 0.09	-0.063	-0.09 ± 0.07
S_7	0.000	0.00 ± 0.09	0.000	0.01 ± 0.06
S_8	0.000	0.00 ± 0.10	0.000	0.00 ± 0.07
$A_{\rm FB}$	0.128	0.14 ± 0.07	0.051	0.06 ± 0.05

that are allowed to vary independently for the S_i^- and S_i^- categories. Namely, these are the yield of the signal and the yield and slope of the combinatorial background, for the $B^0 \to K^{*0}\mu^+\mu^-$ mode, and the yield of the signal, the yield and slope of the combinatorial background, and the yield of the partially reconstructed decays, for the $B^0 \to K^{*0}e^+e^-$ mode.

The expected sensitivity based on the statistics observed in data is estimated from signal-only pseudo-experiments, as described in Section 7.3.1. The counting method approach is validated using both simulation and data (see Section 7.3.2) and is then used to compute the values of the S_i and ΔS_i observables of the $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays, which are given in Section 7.3.3.

7.3.1 Sensitivity studies

The sensitivity of the S_4, S_5, S_7, S_8, and $A_{\rm FB}$ measurements is studied using several ensembles of simulated pseudo-experiments of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays. The $B^0 \to K^{*0}\mu^+\mu^-$ decays are generated according to the SM and NP values of the angular observables shown in Tables 7.7 and 7.8. The NP scenario corresponds to NP contributions to the $C_{9\mu}$ Wilson coefficient only, with $C_{9\mu}^{\rm NP} = -1.4$. This is one of the NP scenarios considered in Section 6.1.2. The $B^0 \to K^{*0}e^+e^-$ decays are instead generated according to the SM values reported in Table 7.9. It is important to note that no uncertainty is associated to the values used in the generation of the pseudo-experiments. However, this can be non-negligible even for SM predictions, as shown in Fig. 6.2. The $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays are generated in two regions of q^2, following the choice of the low and central q^2 bins adopted in this analysis. A total of 500 pseudo-experiments are generated for each model, decay mode, and q^2 bin. The signal yield of the pseudo-experiments is allowed to fluctuate according

Table 7.9: Values of the S_i observables used in the generation of the $B^0 \to K^{*0} e^+ e^-$ pseudo-experiments, compared to the values measured with the counting method approach.

S_i	Low q^2		Central q^2	
	Generated	Measured	Generated	Measured
S_4	-0.075	-0.11 ± 0.17	0.138	0.19 ± 0.12
S_5	0.208	0.29 ± 0.14	-0.188	-0.26 ± 0.10
S_7	0.000	0.00 ± 0.15	0.000	0.00 ± 0.10
S_8	0.000	-0.01 ± 0.18	0.000	0.01 ± 0.12
$A_{\rm FB}$	0.117	0.12 ± 0.11	-0.050	-0.08 ± 0.08

Table 7.10: Values of the ΔS_i observables used in the generation of the pseudo-experiments according to the SM, compared to the values measured with the counting method approach.

ΔS_i	Low q^2		Central q^2	
	Generated	Measured	Generated	Measured
ΔS_4	-0.010	-0.02 ± 0.20	0.001	0.00 ± 0.14
ΔS_5	0.015	0.03 ± 0.17	-0.002	0.00 ± 0.13
ΔS_7	0.000	-0.01 ± 0.18	0.000	-0.01 ± 0.12
ΔS_8	0.000	-0.01 ± 0.20	0.000	0.00 ± 0.14
$\Delta A_{\rm FB}$	0.006	0.01 ± 0.12	-0.001	0.00 ± 0.10

to a Poisson distribution with the mean corresponding to the signal yield observed in data.

Since the samples of the pseudo-experiments described in this section do not contain any background contribution, it is possible to obtain the values of the S_i observables by simply counting the number of candidates in the S_i^+ and S_i^- categories. The distribution of the results is shown in Figs. 7.10 and 7.12 for the muon and electron modes generated according to the SM and in Fig. 7.11 for the muon mode generated according to the NP scenario. The curve corresponding to the fit with a Gaussian distribution is also shown in the plots. The pulls of these distributions with respect to the generated values are also studied and found to be compatible with a Gaussian distribution of mean equal to zero and width equal to unity.

The sensitivity of the ΔS_4, ΔS_5, ΔS_7, ΔS_8, and $\Delta A_{\rm FB}$ measurements is studied analogously. The results are shown in Figs. 7.14 and 7.15 together with the curve corresponding to the Gaussian fit, for both the SM and NP scenarios. The mean and width values obtained from the Gaussian fit are reported in Tables 7.10 and 7.11 and shown in Fig. 7.13. The presence of NP results in a shift of the central values of the ΔS_5 and $\Delta A_{\rm FB}$ observables in the central q^2 bin. The relative uncertainty expected in this region is of nearly 71% for the ΔS_5 measurement and of 64% for $\Delta A_{\rm FB}$. This estimate is based on signal-only pseudo-experiments, while a more realistic estimate based on both signal and background contributions is discussed in the following.

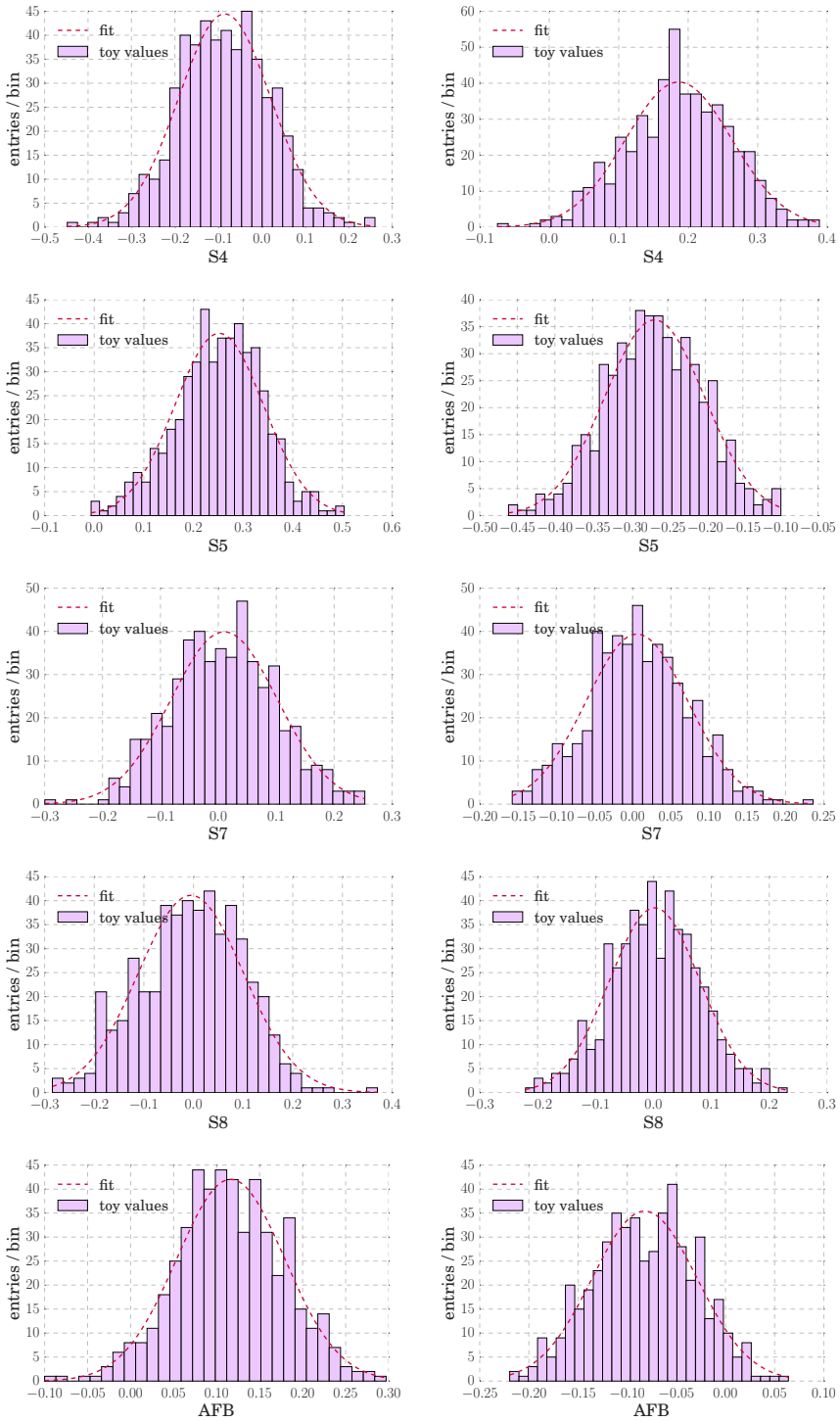


Figure 7.10: Computed values of the S_4 (first row), S_5 (second row), S_7 (third row), S_8 (fourth row), and $A_{\rm FB}$ (fifth row) observables, obtained for the $B^0 \to K^{*0}\mu^+\mu^-$ pseudo-experiments generated according to the SM in the low (left) and central (right) q^2 bins. The results of the fit with a Gaussian distribution are shown by the dashed line in red.

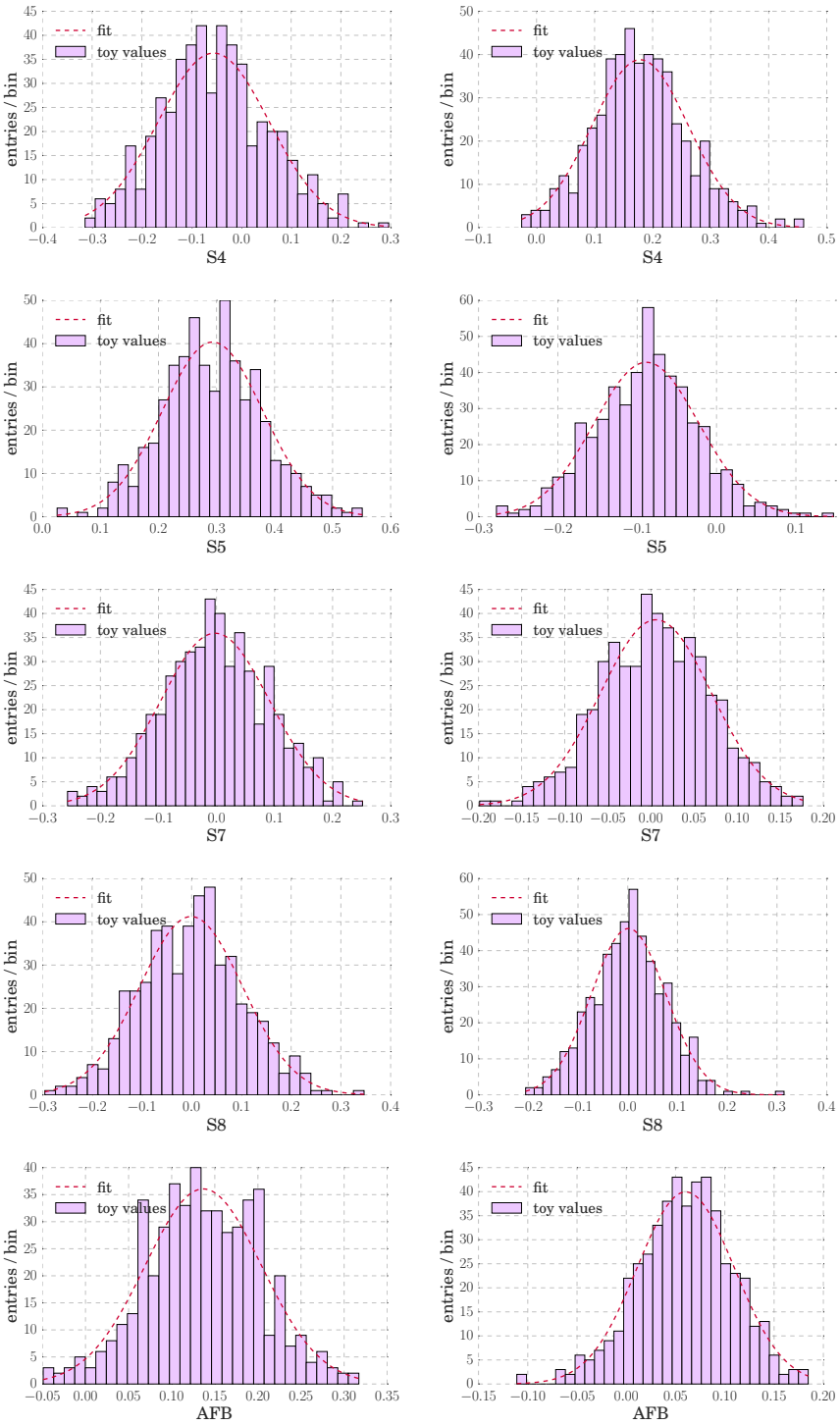


Figure 7.11: Computed values of the S_4 (first row), S_5 (second row), S_7 (third row), S_8 (fourth row), and $A_{\rm FB}$ (fifth row) observables, obtained for the $B^0 \to K^{*0}\mu^+\mu^-$ pseudo-experiments generated according to the NP scenario in the low (left) and central (right) q^2 bins. The results of the fit with a Gaussian distribution are shown by the dashed line in red.

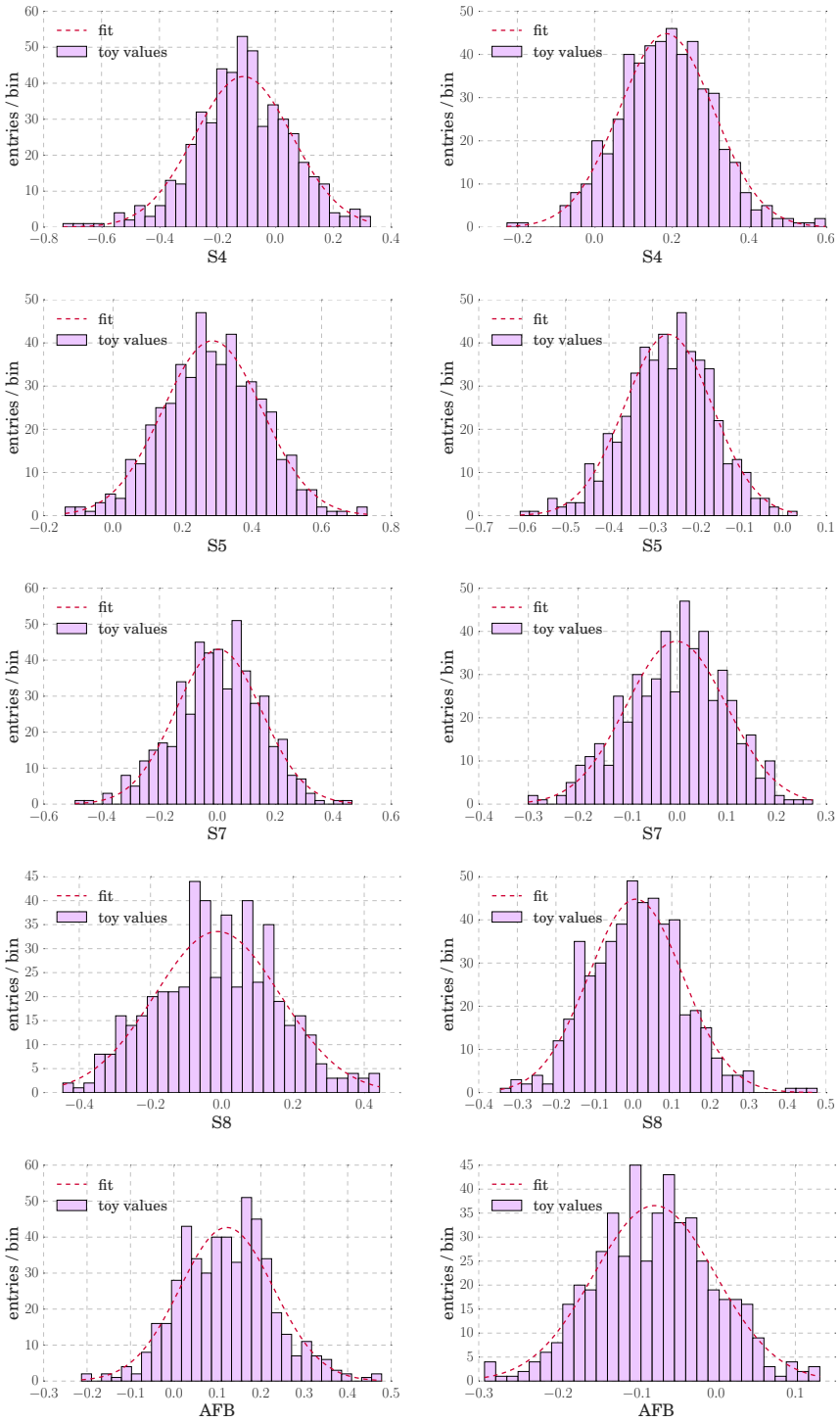


Figure 7.12: Computed values of the S_4 (first row), S_5 (second row), S_7 (third row), S_8 (fourth row), and $A_{\rm FB}$ (fifth row) observables, obtained for the $B^0 \to K^{*0} e^+ e^-$ pseudo-experiments generated according to the SM in the low (left) and central (right) q^2 bins. The results of the fit with a Gaussian distribution are shown by the dashed line in red.

Table 7.11: Values of the ΔS_i observables used in the generation of the pseudo-experiments according to the NP, compared to the values measured with the counting method approach.

ΔS_i	Low q^2		Central q^2	
	Generated	Measured	Generated	Measured
ΔS_4	-0.039	-0.05 ± 0.20	0.001	0.01 ± 0.14
ΔS_5	-0.034	-0.01 ± 0.17	-0.125	-0.17 ± 0.12
ΔS_7	0.000	0.00 ± 0.18	0.000	-0.01 ± 0.12
ΔS_8	0.000	0.00 ± 0.20	0.000	0.00 ± 0.15
$\Delta A_{\rm FB}$	-0.011	-0.02 ± 0.12	-0.101	-0.14 ± 0.09

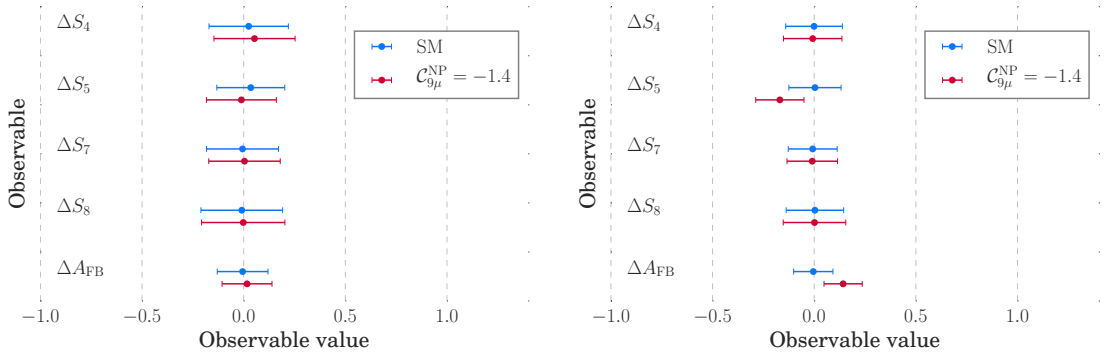


Figure 7.13: Expected sensitivity of the ΔS_4, ΔS_5, ΔS_7, ΔS_8, and $\Delta A_{\rm FB}$ measurements, as obtained from pseudo-experiments of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays in the low (left) and central (right) q^2 bins. The sensitivity to the SM is shown in blue, while the sensitivity to a NP scenario with $\mathcal{C}_{9\mu}^{\rm NP} = -1.4$ is shown in red.

7.3.2 Method validation

Validation with LHCb simulation

The values of the S_i observables are computed using the official LHCb simulation of the rare modes $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ and the resonant modes $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ and $B^0 \to K^{*0}J/\psi$ $(\to \mu^+\mu^-)$. This is done at two different stages of the simulation, referred to as *generation-level* and *reconstruction-level*. The first consists of a simple simulation of the decays without taking the response of the LHCb detector into account, while the second consists of a full simulation, which includes also the reconstruction, trigger, and selection criteria discussed in Chapter 6. As described in Section 6.8.2, the acceptance of the detector generates an asymmetry in the distribution of the decay angles, which can affect differently the S_i^+ and S_i^- categories. As a consequence, the values of the S_i observables might differ depending on whether they are computed from generated or reconstructed candidates. If the reconstruction-level simulation is corrected for the acceptance of the detector, however, the computation of the S_i observables should held compatible results between generated and reconstructed samples.

The values of the S_i observables obtained for the $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ simulated candidates are shown in Table 7.12 for both the generation-level and reconstruction-level samples. For the latter, in particular, two different values are computed. The first is obtained by simply counting the number

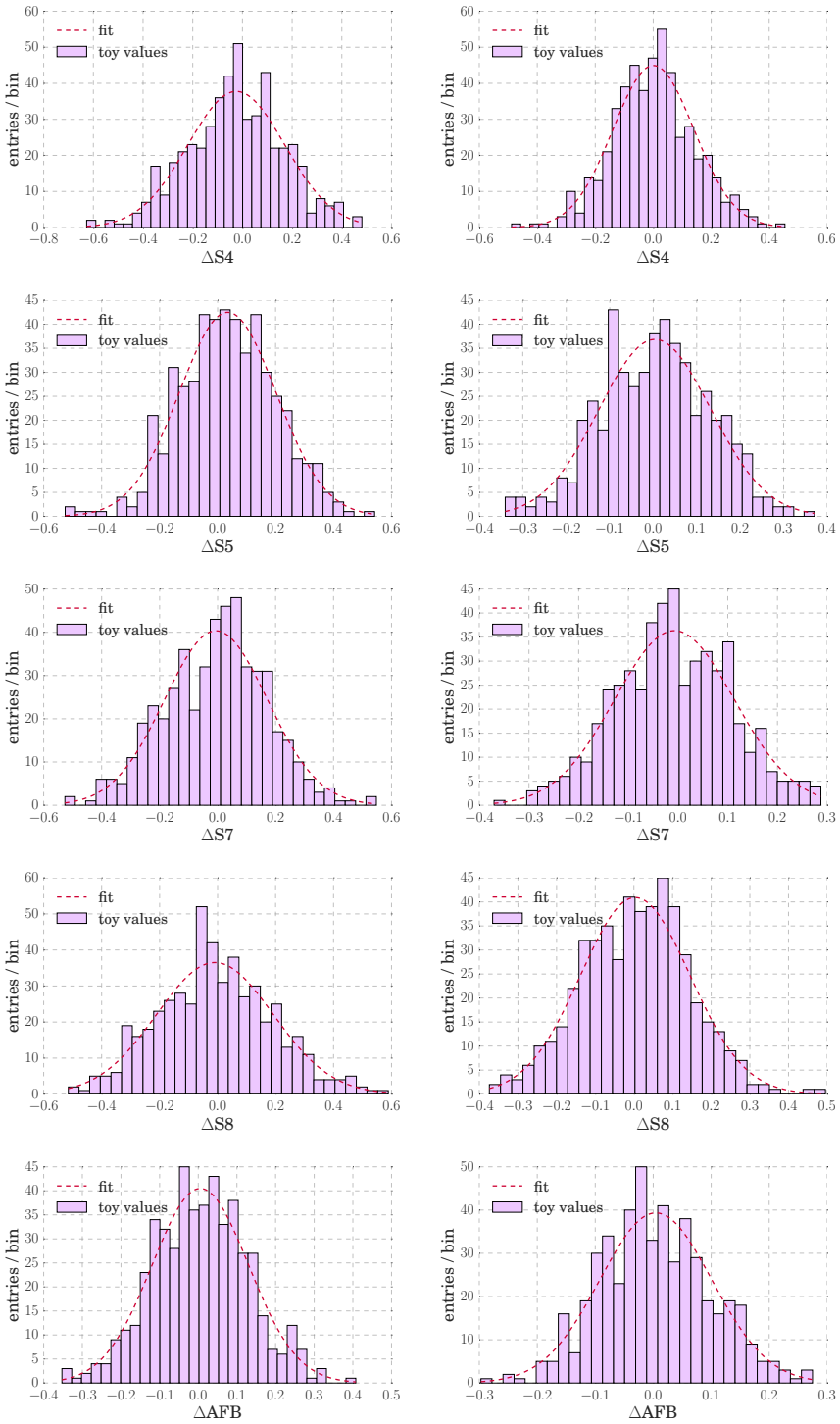


Figure 7.14: Computed values of the ΔS_4 (first row), ΔS_5 (second row), ΔS_7 (third row), ΔS_8 (fourth row), and $\Delta A_{\rm FB}$ (fifth row) observables, obtained for the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ pseudo-experiments generated according to the SM in the low (left) and central (right) q^2 bins. The results of the fit with a Gaussian distribution are shown by the dashed line in red.

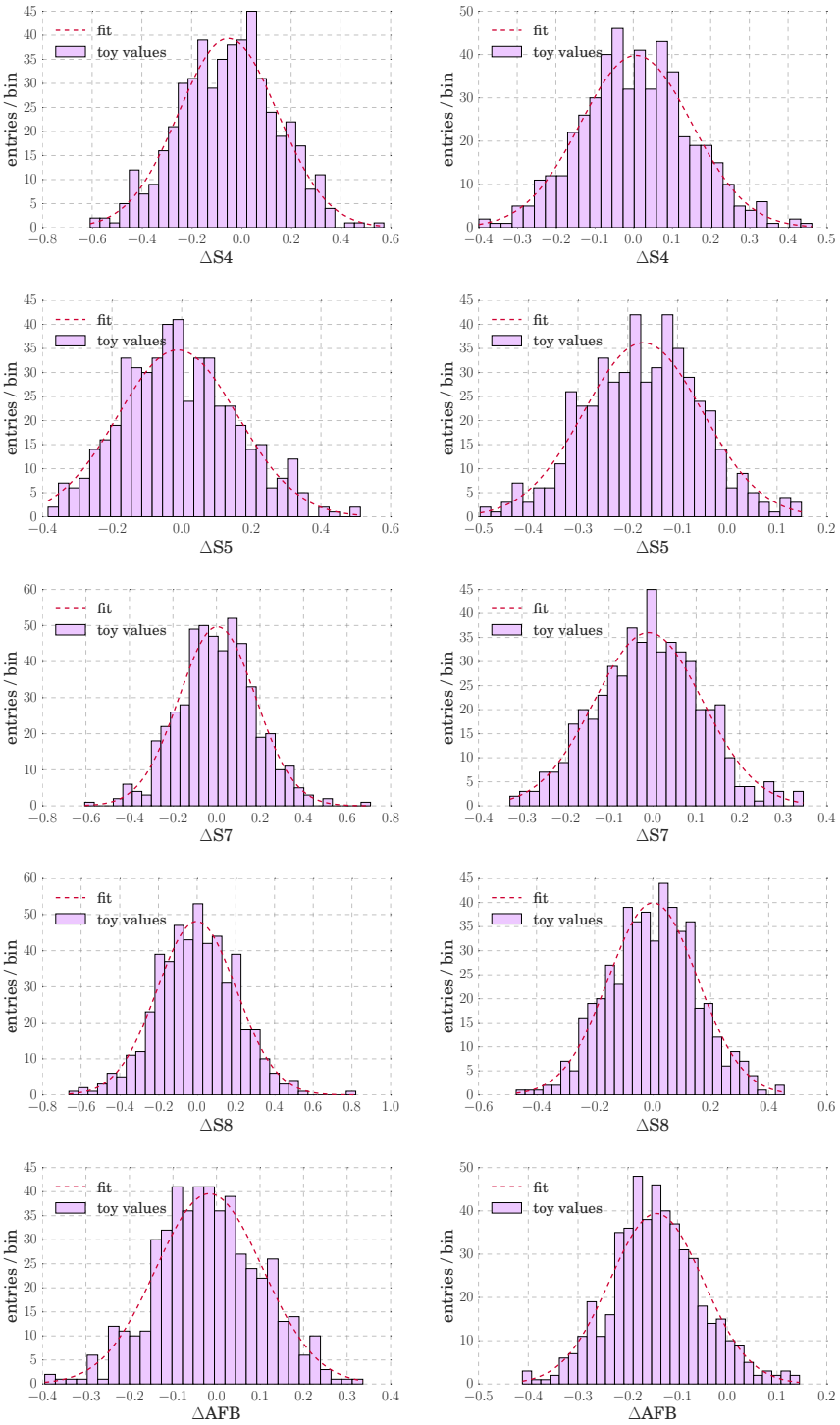


Figure 7.15: Computed values of the ΔS_4 (first row), ΔS_5 (second row), ΔS_7 (third row), ΔS_8 (fourth row), and $\Delta A_{\rm FB}$ (fifth row) observables, obtained for the $B^0 \to K^{*0}\mu^+\mu^-$ pseudo-experiments generated according to the NP scenario and the $B^0 \to K^{*0}e^+e^-$ pseudo-experiments generated according to the SM in the low (left) and central (right) q^2 bins. The results of the fit with a Gaussian distribution are shown by the dashed line in red.

Table 7.12: Comparison between generation-level and reconstruction-level S_i observables, obtained from the $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) simulation sample. Two values are computed at reconstruction-level, with and without the acceptance correction. The reconstruction-level values that show a deviation larger than two standard deviations from the corresponding generation-level values are highlighted in red.

S_i	Trigger	Gen	Rec (without acc)	Rec (with acc)
S_4	LOE	-0.2175 ± 0.0005	-0.260 ± 0.005	-0.222 ± 0.005
	LOH		-0.261 ± 0.014	-0.218 ± 0.015
	LOTIS		-0.278 ± 0.009	-0.219 ± 0.010
S_5	LOE	-0.004 ± 0.007	0.005 ± 0.004	-0.001 ± 0.004
	LOH		0.002 ± 0.012	0.005 ± 0.013
	LOTIS		-0.004 ± 0.007	-0.002 ± 0.009
S_7	LOE	0.012 ± 0.007	0.004 ± 0.004	-0.002 ± 0.004
	LOH		-0.024 ± 0.012	-0.012 ± 0.013
	LOTIS		0.012 ± 0.007	0.005 ± 0.009
S_8	LOE	-0.011 ± 0.009	0.001 ± 0.005	0.001 ± 0.005
	LOH		-0.011 ± 0.014	-0.009 ± 0.015
	LOTIS		-0.011 ± 0.009	0.007 ± 0.010
$A_{\rm FB}$	LOE	-0.001 ± 0.006	-0.0011 ± 0.0029	-0.0016 ± 0.0033
	LOH		-0.002 ± 0.009	0.007 ± 0.010
	LOTIS		-0.001 ± 0.006	-0.002 ± 0.007

of candidates belonging to the S_i^+ and S_i^- categories. The second is evaluated after the simulation sample is corrected for the acceptance effects described in Section 6.8.2 and is computed using the sum of the acceptance weights. The mismatch between generation-level and reconstruction-level values is particularly strong in the case of the S_4 observable. If the acceptance correction is taken into account, a good agreement between generation-level and reconstruction-level values is retrieved for all S_i observables under investigation. Similar results are observed for the other decay modes. When measuring the S_i and ΔS_i observables in data, the effect of the acceptance is taken into account in the efficiency terms.

Validation with data

The resonant mode $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) is used to cross-check the measurement of the S_i observables in data.

The computed values of the S_i observables of the $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) data candidates are reported in Table 7.13. For simplicity, only the results of the LOE trigger category of the electron mode are shown in the table. Similar results are obtained for the other trigger categories. The obtained values are compatible with previous measurements at LHCb [134].

Expected sensitivity and validation with pseudo-experiments

The expected sensitivity of the S_i measurement in the electron mode is evaluated by generating 1000 pseudo-experiments of $B^0 \to K^{*0} e^+e^-$ decays in the low and central q^2 bins. When generating the pseudo-experiments, the candidates are sampled from the distributions obtained from the simultaneous fit to the S_i^+ and S_i^- categories in data. Therefore, each pseudo-experiment is a

Table 7.13: Results of the fit to the $B^0 \to K^{*0} J/\psi$ ($\to e^+e^-$) data candidates in the L0E trigger category.

S_i	$\mathcal{N}^+_{\rm sig}$	$\mathcal{N}^+_{\rm bkg}$	$\mathcal{N}^-_{\rm sig}$	$\mathcal{N}^-_{\rm bkg}$	S_i value
S_4	14 432±140	42±60	20 382±140	59±20	−0.234±0.019
S_5	17 353±240	52±280	17 468±120	42±15	−0.009±0.014
S_7	17 362±130	33±16	17 457±130	63±25	0.012±0.014
S_8	17 561±130	31±14	17 233±150	90±70	−0.012±0.017
$A_{\rm FB}$	17 308±170	56±100	17 505±130	46±17	−0.004±0.011

Table 7.14: Results of the Gaussian fit to the distribution of the S_i observables computed from the 1000 $B^0 \to K^{*0} e^+ e^-$ pseudo-experiments with signal and background in the low and central q^2 bins.

S_i	Low q^2	Central q^2
S_4	0.37±0.32	−0.06±0.24
S_5	−0.03±0.27	−0.04±0.21
S_7	0.16±0.26	0.04±0.20
S_8	−0.34±0.33	−0.10±0.24
$A_{\rm FB}$	−0.16±0.20	0.08±0.15

mixture of signal and background. This allows to obtain a more realistic estimate of the sensitivity compared to the signal-only scenario discussed in Section 7.3.1. The total yield of each pseudo-experiment is allowed to fluctuate according to a Poisson distribution with the mean fixed to the total yield measured in data. The candidates are then fitted using the same model of the fit to the data. The fitted signal yields are used to compute the S_i observables of each pseudo-experiment. The resulting distributions and their fit with a Gaussian distribution are shown in Fig. 7.16 for both low and central q^2 bins. The mean and width of the Gaussian fit are summarised for completeness in Table 7.14. The statistical uncertainty reported in the table corresponds to the expected sensitivity of the S_i measurement in data. As expected, these values are larger compared to those of the signal-only scenario.

The same pseudo-experiments described above are used to investigate the stability and unbiasedness of the mass fits. This is done by studying the distributions of the fit results and their pulls with respect to the values used in the generation of the pseudo-experiments. The fit results for the $B^0 \to K^{*0} e^+ e^-$ candidates in the low q^2 bin are shown in Fig. 7.17, while those obtained for the central q^2 bin are shown in Fig. 7.18. For each q^2 bin, the distributions of the fitted yields of signal and background are shown. The pulls of the signal yields with respect to the values used in the generation of the pseudo-experiments are shown in Fig. 7.19. The distributions in Figs. 7.17, 7.18, and 7.19 refer to the S_5 observable, but similar results are obtained for the other observables. The fitted yields and their pulls are fitted with a Gaussian distribution in order to determine whether the fit is unbiased. As expected, the pulls are compatible with a standard normal distribution. For the low q^2 bin, where the statistics is rather limited, a small bias is observed in the background yields. It would be useful to repeat the same study using Minos [135] in the fit of the pseudo-experiments, since this is expected to provide a better estimate of the best-fit parameters. If no change is observed, one possible way to address this issue would be to apply a selection based on the HOP variable to the

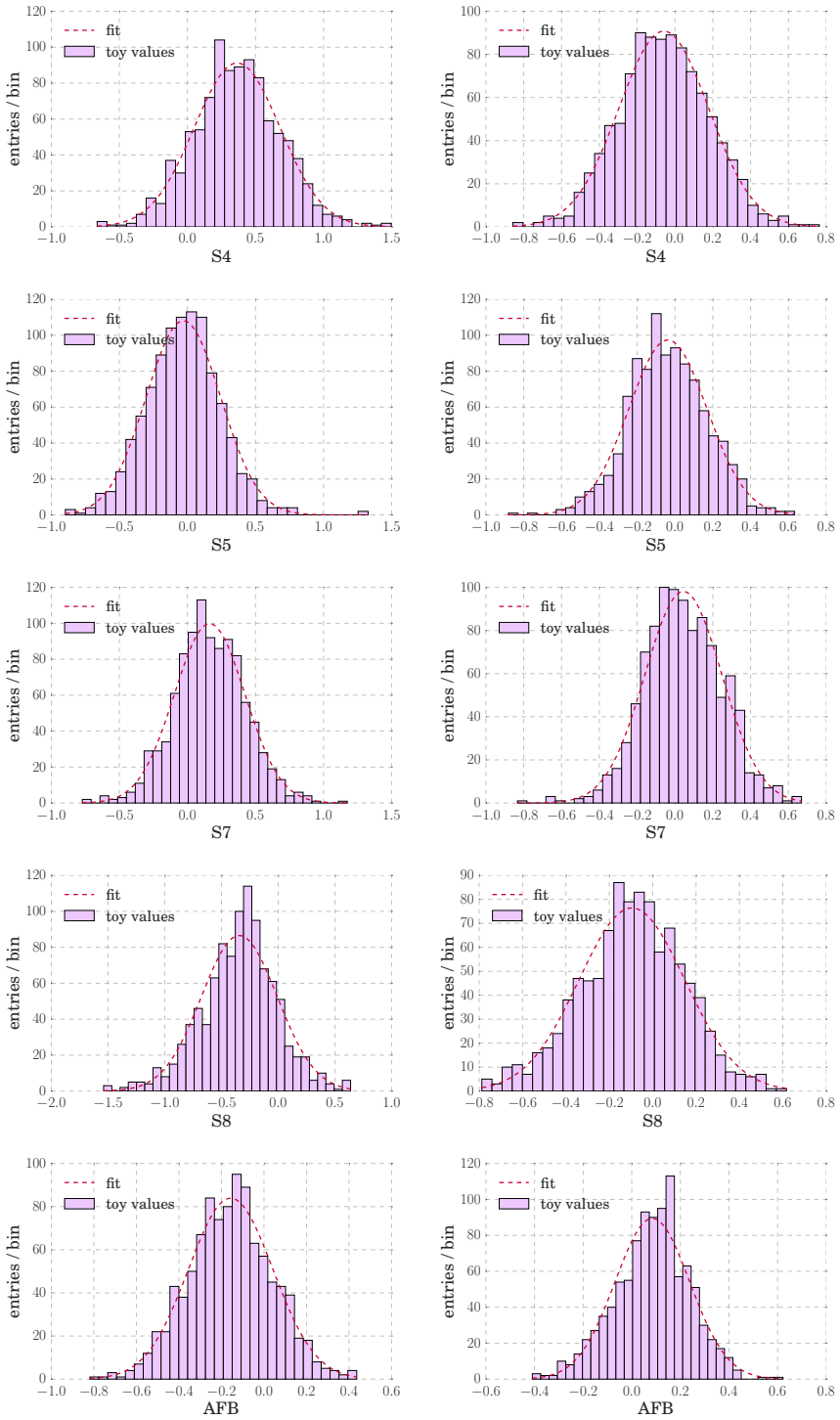


Figure 7.16: Computed values of the S_4 (first row), S_5 (second row), S_7 (third row), S_8 (fourth row), and $A_{\rm FB}$ (fifth row) observables, obtained for the $B^0 \to K^{*0} e^+ e^-$ 1000 pseudo-experiments of signal and background in the low (left) and central (right) q^2 bins. The results of the fit with a Gaussian distribution are shown by the dashed line in red.

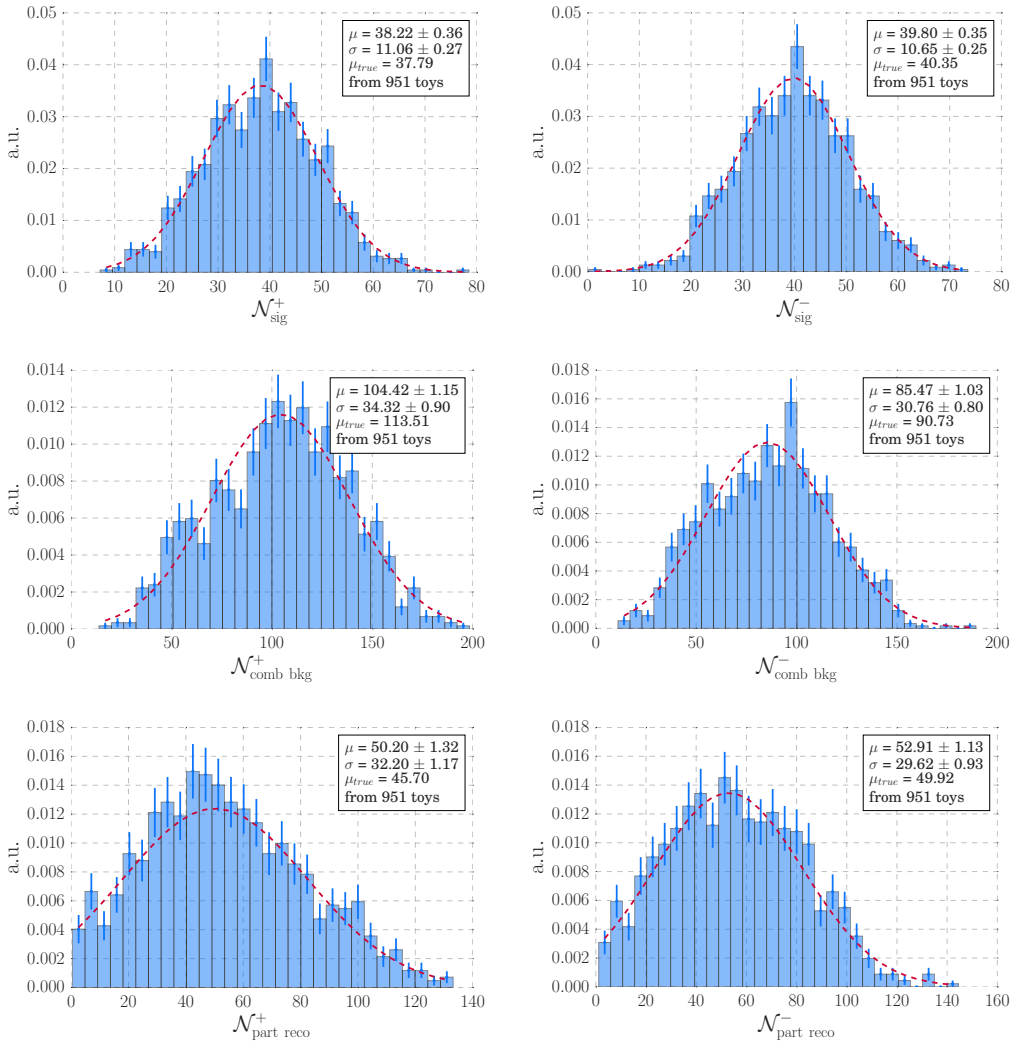


Figure 7.17: Distribution of the fitted yields of signal (top), combinatorial background (middle), and partially reconstructed background (bottom) obtained from the $B^0 \to K^{*0}e^+e^-$ 1000 pseudo-experiments of signal and background, for the low q^2 bin. The curve corresponding to the Gaussian fit and the parameters of the curve are also shown.

$B^0 \to K^{*0}e^+e^-$ candidates in the low q^2 bin and repeat the mass fit. This additional requirement is expected to lower the contamination from partially reconstructed decays in the mass window where the fit is performed.

7.3.3 Results

Measurement of the S_i observables

The signal and background yields obtained from the fit to the $B^0 \to K^{*0}\mu^+\mu^-$ candidates in data are reported in Tables 7.15 and 7.16 for the low and the central q^2 bin, respectively. The results of the measurement of the S_i observables in the low q^2 bin are shown in Table 7.17, while those in the central q^2 bin are shown in Table 7.18. Two different choices of the low q^2 bin are considered: the

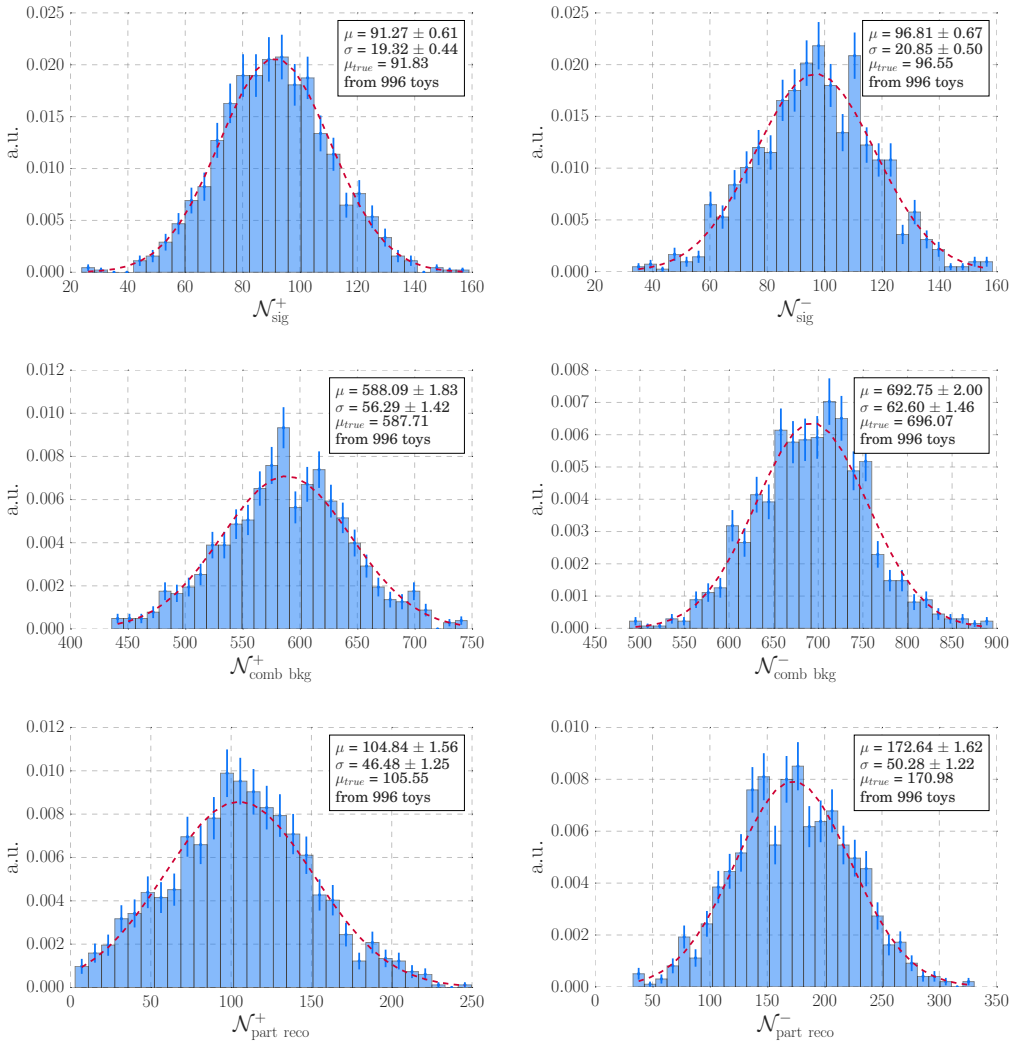


Figure 7.18: Distribution of the fitted yields of signal (top), combinatorial background (middle), and partially reconstructed background (bottom) obtained from the $B^0 \to K^{*0}e^+e^-$ 1000 pseudo-experiments of signal and background, for the central q^2 bin. The curve corresponding to the Gaussian fit and the parameters of the curve are also shown.

baseline q^2 bin of the measurement described in this thesis, from $0.1\,\text{GeV}^2/c^4$ to $1.1\,\text{GeV}^2/c^4$, and an alternative q^2 bin from $0.045\,\text{GeV}^2/c^4$ to $1.1\,\text{GeV}^2/c^4$. The S_i observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay are computed also in the q^2 bin from $1.1\,\text{GeV}^2/c^4$ to $6.0\,\text{GeV}^2/c^4$, in order to compare the results of the counting method procedure with the results of the LHCb measurement from Ref. [35]. The results are summarised in Table 7.18, where they are compared to the values measured in Ref. [35]. The measured values are in agreement with those reported in Ref. [35] for all angular observables of interest.

The results in the nominal q^2 bins of this measurement are compared to the predictions of the SM and the NP scenario described in Section 7.3.1. The agreement between the measured values and the predictions is displayed in Fig. 7.20, where the error bars of the experimental results are statistical only and the error bars of the predictions correspond to the signal-only expected sensitivity. It is

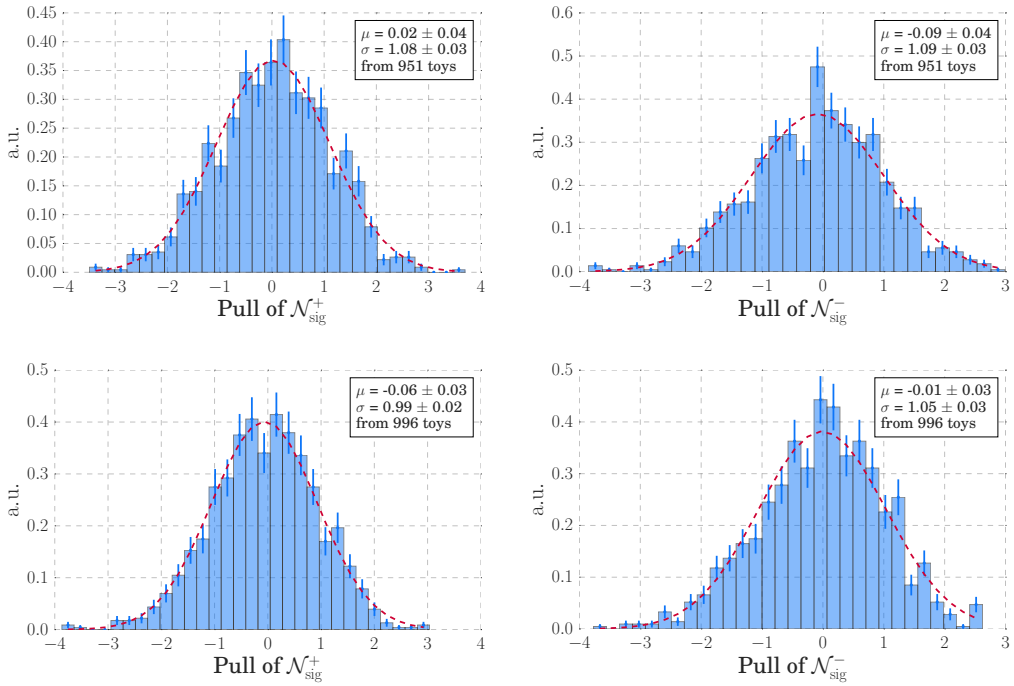


Figure 7.19: Pulls of the fitted signal yields obtained from the $B^0 \to K^{*0}e^+e^-$ 1000 pseudo-experiments of signal and background, for the low (top) and central (right) q^2 bins.

important to note that the sign of the predictions for the S_4 and $A_{\rm FB}$ observables is the opposite with respect to what shown in Section 7.3.1. This is needed in order to correctly compare the predictions and the experimental results [136], due to the different conventions adopted by LHCb and the theory community in the definition of the decay angles. The results are dominated by the statistical uncertainties and are compatible with both SM and NP predictions.

The impact of the combined selection in q^2 and constrained q^2 is evaluated by determining the S_i observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay with and without the requirement on the constrained q^2. The results corresponding to the two selections in q^2 are compatible with each other within the uncertainties for all S_i observables.

The fit results for the $B^0 \to K^{*0}e^+e^-$ decay are reported in Tables 7.19 and 7.20 for the low and the central q^2 bin, respectively. The results of the S_i measurement are reported in Table 7.21 for the low q^2 bin and in Table 7.22 for the central q^2 bin. The statistical uncertainty of the S_i measurements is consistent with the sensitivity obtained from the realistic pseudo-experiments of Section 7.3.2. The results are dominated by the statistical uncertainty and are compatible with the SM predictions, as shown in Fig. 7.21.

Measurement of the ΔS_i observables

The results of the measurement of the ΔS_i observables between the $B^0 \to K^{*0}\mu^+\mu^-$ and the $B^0 \to K^{*0}e^+e^-$ decay are reported in Table 7.23 for the low q^2 bin and in Table 7.24 for the central q^2 bin. The expected sensitivity based on the studies of Section 7.3.1 is also shown for comparison. It is important to note that the uncertainty of the results is statistical only and is

Table 7.15: Results of the fit to the $B^0 \to K^{*0}\mu^+\mu^-$ data candidates in the low q^2 bin from $0.1\,\text{GeV}^2/c^4$ to $1.1\,\text{GeV}^2/c^4$, obtained from the samples with the selection on the constrained q^2.

S_i	$\mathcal{N}^+_{\text{sig}}$	$\mathcal{N}^+_{\text{bkg}}$	$\mathcal{N}^-_{\text{sig}}$	$\mathcal{N}^-_{\text{bkg}}$
S_4	106±11	21±6	103±11	17±6
S_5	134±13	14±6	77±10	22±6
S_7	103±11	11±5	105±11	28±7
S_8	105±11	18±6	104±11	20±7
A_{FB}	103±11	13±6	108±11	23±6

Table 7.16: Results of the fit to the $B^0 \to K^{*0}\mu^+\mu^-$ data candidates in the central q^2 bin from $1.1\,\text{GeV}^2/c^4$ to $7.0\,\text{GeV}^2/c^4$, obtained from the samples with the selection on the constrained q^2.

S_i	$\mathcal{N}^+_{\text{sig}}$	$\mathcal{N}^+_{\text{bkg}}$	$\mathcal{N}^-_{\text{sig}}$	$\mathcal{N}^-_{\text{bkg}}$
S_4	181±15	78±11	241±17	85±12
S_5	179±15	82±11	244±17	80±12
S_7	212±16	88±12	210±16	75±11
S_8	219±17	81±12	203±16	81±12
A_{FB}	212±16	83±12	210±16	80±12

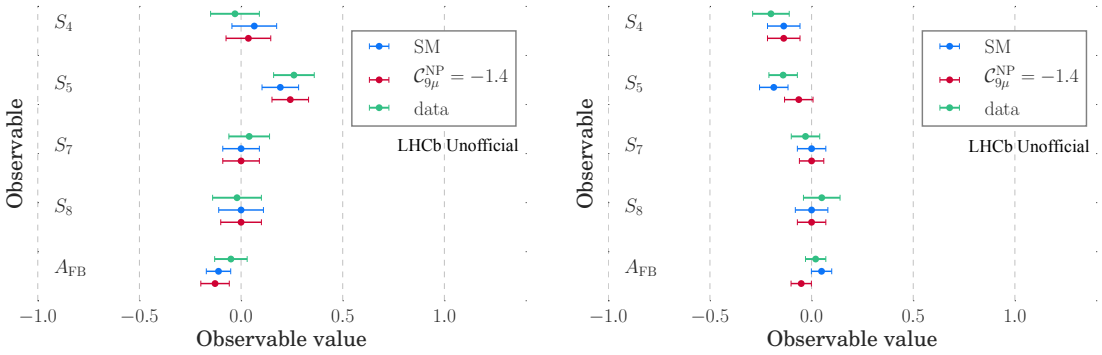


Figure 7.20: Results of the measurement of the S_4, S_5, S_7, S_8, and A_{FB} observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay in the low (left) and central (right) q^2 bins, compared to the expected sensitivity described in Fig. 7.13. The data points are shown in green.

larger than the expected sensitivity due to the fact that only the signal was taken into account in the pseudo-experiments. The results are compatible with both SM and NP predictions within the uncertainties for all angular observables of interest. The agreement with the results of the pseudo-experiment studies discussed in Section 7.3.1 is shown in Fig. 7.22, where the measurement is represented by the green error bars.

Table 7.17: Results of the measurement of the S_i observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay in the q^2 bin from $0.045\,\text{GeV}^2/c^4$ to $1.1\,\text{GeV}^2/c^4$ and in the q^2 bin from $0.1\,\text{GeV}^2/c^4$ to $1.1\,\text{GeV}^2/c^4$.

S_i	$[0.045, 1.1]\,\text{GeV}^2/c^4$	$[0.1, 1.1]\,\text{GeV}^2/c^4$
S_4	0.03±0.10	−0.03±0.12
S_5	0.23±0.09	0.26±0.10
S_7	0.03±0.09	0.04±0.10
S_8	−0.08±0.10	−0.02±0.12
A_{FB}	−0.07±0.07	−0.05±0.08

Table 7.18: Results of the measurement of the S_i observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay in the q^2 bin from $1.1\,\text{GeV}^2/c^4$ to $7.0\ \text{GeV}^2/c^4$ and in the q^2 bin from $1.1\,\text{GeV}^2/c^4$ to $6.0\,\text{GeV}^2/c^4$, compared to the LHCb measurement reported in Ref. [35].

S_i	$[1.1, 7.0]\,\text{GeV}^2/c^4$	$[1.1, 6.0]\,\text{GeV}^2/c^4$	LHCb measurement
S_4	−0.20±0.09	−0.24±0.10	$-0.155^{+0.057}_{-0.056} \pm 0.004$
S_5	−0.14±0.00	−0.04±0.08	$-0.023^{+0.050}_{-0.049} \pm 0.005$
S_7	−0.03±0.07	−0.05±0.08	$-0.077^{+0.050}_{-0.049} \pm 0.006$
S_8	0.05±0.09	0.11±0.10	$0.028^{+0.058}_{-0.057} \pm 0.008$
A_{FB}	0.02±0.05	0.00±0.06	$-0.075^{+0.032}_{-0.034} \pm 0.007$

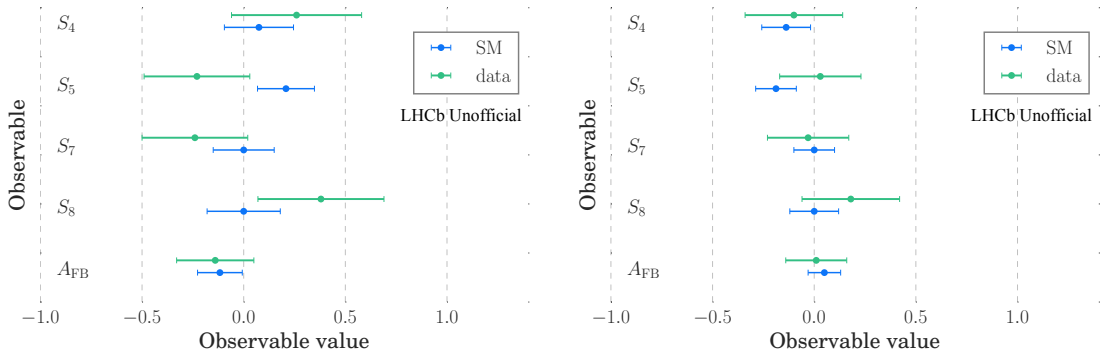


Figure 7.21: Results of the measurement of the S_4, S_5, S_7, S_8, and A_{FB} observables of the $B^0 \to K^{*0}e^+e^-$ decay in the low (left) and central (right) q^2 bins, compared to the expected sensitivity described in Fig. 7.13. The data points are shown in green.

7.4 Systematic uncertainties

The evaluation of the systematic uncertainties for the measurement of the S_i and ΔS_i observables is still ongoing. This section describes the contributions that are being taken into account and the strategy to assign a systematic uncertainty to the final measurement. Two classes of systematic uncertainties are described in the following. The first addresses the systematic uncertainties associated to parameters constrained in the fit to values obtained from simulation and to the model used for the nominal fit. The second describes the systematic uncertainties related to the corrections applied to the simulated candidates.

Table 7.19: Results of the fit to the $B^0 \to K^{*0} e^+ e^-$ data candidates in the low q^2 bin from $0.1\,\mathrm{GeV}^2/c^4$ to $1.1\,\mathrm{GeV}^2/c^4$, obtained from the samples with the selection on the constrained q^2.

S_i	$\mathcal{N}^+_{\text{sig}}$	$\mathcal{N}^+_{\text{comb bkg}}$	$\mathcal{N}^+_{\text{part reco}}$	$\mathcal{N}^-_{\text{sig}}$	$\mathcal{N}^-_{\text{comb bkg}}$	$\mathcal{N}^-_{\text{part reco}}$
S_4	46±11	94±30	40±25	32±11	106±32	60±27
S_5	37±11	107±31	54±26	42±11	89±34	49±28
S_7	32±10	118±30	37±25	48±11	74±37	68±32
S_8	48±10	63±30	77±28	29±11	135±30	26±25
A_{FB}	34±9	108±32	88±29	44±12	89±26	16±21

Table 7.20: Results of the fit to the $B^0 \to K^{*0} e^+ e^-$ data candidates in the central q^2 bin from $1.1\,\mathrm{GeV}^2/c^4$ to $7.0\,\mathrm{GeV}^2/c^4$, obtained from the samples with the selection on the constrained q^2.

S_i	$\mathcal{N}^+_{\text{sig}}$	$\mathcal{N}^+_{\text{bkg}}$	$\mathcal{N}^+_{\text{part reco}}$	$\mathcal{N}^-_{\text{sig}}$	$\mathcal{N}^-_{\text{bkg}}$	$\mathcal{N}^-_{\text{part reco}}$
S_4	82±18	555±60	116±50	100±21	714±70	217±60
S_5	86±20	582±60	118±50	95±19	687±70	214±60
S_7	89±19	646±70	178±50	93±20	621±60	156±50
S_8	100±19	622±70	183±50	80±19	646±60	150±50
A_{FB}	94±22	1016±80	228±60	87±16	243±50	114±38

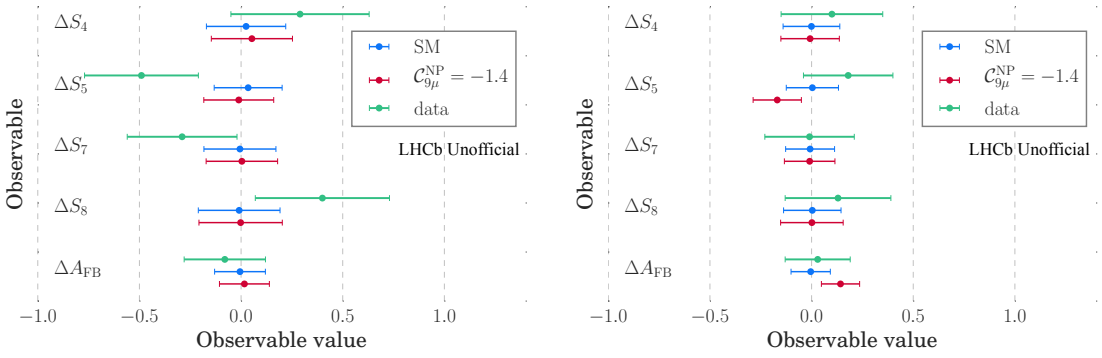


Figure 7.22: Results of the measurement of the ΔS_4, ΔS_5, ΔS_7, ΔS_8, and ΔA_{FB} observables of the $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0} e^+ e^-$ decays in the low (left) and central (right) q^2 bins, compared to the expected sensitivity described in Fig. 7.13. The data points are shown in green.

7.4.1 Mass model and fit strategy

Systematic uncertainties are assigned according to one of the two procedures:

- An ensemble of pseudo-experiments is generated according to an alternative model and is then fitted according to the nominal strategy;
- A different method is used to compute the observables of interest directly in data.

The first procedure can be used to probe different models for the signal and background components in the mass fit, for example by replacing the DCB of the signal with a CB or the

Table 7.21: Results of the S_i measurement for the $B^0 \to K^{*0}e^+e^-$ decay in the low q^2 bin. The relative uncertainty of the measurement and the expected sensitivity based on pseudo-experiments are also shown.

S_i	Observable value	Relative uncertainty	Expected sensitivity
S_4	0.26±0.32	1.23	0.32
S_5	−0.23±0.26	1.13	0.27
S_7	−0.24±0.26	1.08	0.26
S_8	0.38±0.31	0.82	0.33
$A_{\rm FB}$	−0.14±0.19	1.36	0.20

Table 7.22: Results of the S_i measurement for the $B^0 \to K^{*0}e^+e^-$ decay in the central q^2 bin. The relative uncertainty of the measurement and the expected sensitivity based on pseudo-experiments are also shown.

S_i	Observable value	Relative uncertainty	Expected sensitivity
S_4	−0.10±0.24	2.40	0.24
S_5	0.03±0.20	6.67	0.21
S_7	−0.03±0.20	6.67	0.20
S_8	0.18±0.24	1.33	0.24
$A_{\rm FB}$	0.01±0.15	15.00	0.15

exponential distribution of the combinatorial background with a linear distribution or a Chebyshev polynomial.

In the mass fit to the $B^0 \to K^{*0}e^+e^-$ data candidates, the shift of the mean and the scale factor of the width of the mass template are obtained from a fit to the PV-constrained B^0 invariant mass distribution of the $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ data candidates, under the implicit assumption that these parameters have the same values for the rare and the resonant mode. A systematic uncertainty to account for this choice can be determined by applying a Gaussian constraint on these two parameters instead of fixing them in the mass fit. Another systematic uncertainty can be assigned to the fact that the template for the partially reconstructed decays is obtained from simulation, in which a $2:1$ combination of K_1^+ and $K_2^*(1430)^+$ decays is used. Since the invariant mass differs from one K^* resonance to another, a systematic uncertainty can be assigned by considering different proportions between K_1^+ and $K_2^*(1430)^+$. For each fraction, pseudo-experiments with the statistics observed in data can be generated and then fitted using the default mass model. Alternatively, the template for the partially reconstructed decays can be obtained directly from data, by applying the *sPlot* technique [127] to the $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ sample. In addition, a systematic uncertainty due to the mass template of the partially reconstructed decays and of the J/ψ decays leaking to the central q^2 region can be determined by modifying the kernel parameter of the non-parametric PDFs. In the central q^2 bin, the yield of $B^0 \to K^{*0}J/\psi$ $(\to e^+e^-)$ decays leaking to the central q^2 region is partially obtained from simulation. A systematic uncertainty can be assigned by leaving this value free to vary in the mass fit.

Another systematic uncertainty can be assigned due to the fact that in the fit to the data some parameters are fixed to the values obtained from the simulation. One possible way to assign a systematic uncertainty would be to generate pseudo-experiments with different values of these

Table 7.23: Results of the measurement of the ΔS_i observables in the low q^2 bin. The expected sensitivity and the SM and NP predictions are also shown. The predictions for S_4 and $A_{\rm FB}$ have the opposite sign with respect to the values reported in Section 7.3.1 in order to correctly compare the experimental results to the predicted values.

ΔS_i	Measurement	SM	NP	Expected sensitivity
ΔS_4	0.29 ± 0.34	0.010	0.039	0.20
ΔS_5	-0.49 ± 0.28	0.015	-0.034	0.17
ΔS_7	-0.29 ± 0.27	0.000	0.000	0.18
ΔS_8	0.40 ± 0.33	0.000	0.000	0.20
$\Delta A_{\rm FB}$	-0.08 ± 0.20	-0.006	0.011	0.12

Table 7.24: Results of the measurement of the ΔS_i observables in the central q^2 bin. The expected sensitivity and the SM and NP predictions are also shown. The predictions for S_4 and $A_{\rm FB}$ have the opposite sign with respect to the values reported in Section 7.3.1 in order to correctly compare the experimental results to the predicted values.

ΔS_i	Measurement	SM	NP	Expected sensitivity
ΔS_4	0.10 ± 0.25	-0.001	-0.001	0.20
ΔS_5	0.18 ± 0.22	-0.002	-0.125	0.17
ΔS_7	-0.01 ± 0.22	0.000	0.000	0.18
ΔS_8	0.13 ± 0.26	0.000	0.000	0.20
$\Delta A_{\rm FB}$	0.03 ± 0.16	0.001	0.101	0.12

parameters, generated according to their correlation matrix, and fit them with the default mass template. The difference between results of the pseudo-experiments and nominal fit can be fitted with a Gaussian function and a systematic uncertainty can be assigned as the linear sum of the absolute value of the corresponding mean and width.

An additional source of systematic uncertainty comes from the fact that the double semileptonic $B^0 \to D^- (\to K^{*0}\ell^-\overline{\nu}_\ell)\ell^+\nu_\ell$ background is not taken into account in the mass fit and can be determined by adding another component in the fit to the $B^0 \to K^{*0}e^+e^-$ data, based on the studies in Section 6.7.2.

The $\Lambda_b^0 \to pK^- J/\psi\ (\to \ell^+\ell^-)$ and $\Lambda_b^0 \to pK^-\ \ell^+\ell^-$ backgrounds are not included in the mass fit to the $B^0 \to K^{*0} J/\psi\ (\to \ell^+\ell^-)$ data candidates. This results in a systematic uncertainty that can be evaluated by adding these components to the mass fit and by comparing the results with those of the nominal fit.

7.4.2 Corrections to simulation

In addition to provide a template for the mass fit to the data samples, the simulation is used to compute the efficiencies in Eq. (6.1) and hence the values of the S_i and ΔS_i observables. Since the simulation is corrected, a systematic uncertainty due to the applied corrections is assigned to the measurement. The three corrections applied to the simulation are described in detail in Section 6.6 and consist in resampling the PID distributions of the simulated candidates and assigning a weight and a correction factor to each event, the first to correct the modelling of the kinematics of the

event and the second to correct the trigger efficiency. The PID correction affects which simulated candidates pass the selection requirements, while the other two corrections modify the sum of the weights that appears in the numerator of the efficiency.

Two systematic uncertainties can be assigned to the PID correction, one due to the limited size, using bootstrapping [137], and another due to the modelling of the data calibration sample used to determine the correction, using different kernel density estimations. The systematic uncertainty due to the multidimensional reweighting can be evaluated by neglecting the corresponding weights in the efficiency calculation. For the trigger efficiency correction, a systematic uncertainty can be assigned by producing new weights based on the largest uncertainties seen in the correction tables and then computing the corresponding efficiency.

It is important to note that, since the S_i observables are computed from an asymmetry between two mass fits, the only effects that are relevant in the evaluation of the systematic uncertainties are those that are not symmetric in S_i^+ and S_i^-. The three corrections to the simulation samples described here should modify in the same way the efficiency of the S_i^+ and S_i^- categories. Therefore, their systematic uncertainties are expected to be negligible compared to the statistical uncertainty of the measurement.

7.5 Future prospects

The analysis described in this thesis is the first attempt at measuring the ΔS_4, ΔS_5, ΔS_7, ΔS_8, and $\Delta A_{\rm FB}$ observables of the $B^0 \to K^{*0} e^+ e^-$ and $B^0 \to K^{*0} \mu^+ \mu^-$ decays at LHCb. This measurement opens the way to further measurements of angular observables, including the Q_i and D_i angular asymmetries, that will allow to probe NP effects involving LFU breaking between electrons and muons.

The measurement shows that it is feasible to increase the upper bound of the central q^2 bin from $6.0\,\text{GeV}^2/c^4$ to $7.0\,\text{GeV}^2/c^4$ by adopting a novel selection based on the constrained q^2. This allows to increase the statistics available and to reduce the contamination from $B^0 \to K^{*0} J/\psi\ (\to e^+ e^-)$ decays leaking in the central q^2 region.

As anticipated in Section 7.4, the evaluation of the systematic uncertainties of the S_i and ΔS_i measurements is still ongoing and is expected to be completed in the next months. However, the systematic uncertainties are expected to be small compared to the large statistical uncertainties. For this reason, it would be advantageous to repeat the same measurement using the full LHC Run I and Run II datasets collected by LHCb. With the increased statistics, a measurement of the angular distribution of the decay products of the $B^0 \to K^{*0} e^+ e^-$ decay similar to what was already done for the muon mode [35] would also become realistic. This would allow to measure the P_i' observables of the electron mode and hence the Q_i angular asymmetries. The Q_i observables are expected to have a larger sensitivity to NP effects compared to the ΔS_i observables, as reported in Ref. [55]. Even if the Q_i angular observables are not directly accessible using the counting method approach, the latter would provide a strong cross-check of the full angular maximum likelihood fit, given its robustness and stability. In addition, further measurements will greatly profit from this first measurement, since the selection requirements, the corrections to the simulation, and the understanding of the experimental challenges related to the reconstruction of electrons in LHCb can be addressed as discussed in this work.

Conclusions

"We do not know what the rules of the game are; all we are allowed to do is to watch the playing. Of course, if we watch long enough, we may eventually catch on to a few of the rules. The rules of the game are what we mean by fundamental physics." – Richard P. Feynman

If one has to take the sentence above as reference, the measurements described in the first part of this thesis focused on how physicists "watch the playing", while the analysis presented in the second part aimed at "catching on to a few of the rules".

The first part of this work presented the test beam measurements performed in view of the upgrade of the LHCb detector. These measurements aimed at characterising the performances of prototype p^+-on-n silicon microstrip sensors for the UT, the tracking detector that will replace the TT after the Long Shutdown II. Irradiated and unirradiated sensors were tested at the SPS at CERN and allowed to understand the performances of two key features of the proposed design. The first feature is the presence of an embedded pitch adapter, that is, a second metallisation layer that allows to adapt the 190 µm pitch of the readout strips to the 80 µm pitch of the input pads of the readout electronics. Two different geometries of the embedded pitch adapter were compared to each other in terms of radiation hardness, efficiency above the full depletion voltage, charge sharing, and cross talk. An efficiency loss localised in the pitch adapter region was observed for the geometry with metal traces and bond pads inside the active area of the sensor. Since this region is small compared to the active area of the sensor, the inefficiency is negligible (at the per mille level) when integrating over the whole sensor. Simulation studies confirmed this effect and showed that the inefficiency can be significantly reduced by increasing the thickness of the silicon oxide layer between readout strips and second metallisation layer. These changes were implemented in the final design of the UT sensors. The second feature is the top-side biasing scheme, that is, the possibility to propagate the high voltage to the backplane through the low-resistivity region of the bulk in the periphery of the sensor. This choice simplifies the routing of the high voltage to the UT modules but can induce a degradation in the signal performance, especially after irradiation. Since high voltage instabilities were observed for the top-side biasing scheme, an improved design was proposed, in which the routing lines are located further from the metal traces of the pitch adapter. The test beam measurements allowed to finalise the design of the UT sensors and showed that the latter can

be operated for the expected run time of the UT.

The second part of this work presented the first measurement of the *CP*-averaged angular observables $A_{\rm FB}$, S_4, S_5, S_7, and S_8 of the $B^0 \to K^{*0} e^+ e^-$ decay with the full LHC Run I dataset collected by the LHCb experiment. The differences between these observables and the corresponding observables of the $B^0 \to K^{*0} \mu^+ \mu^-$ decay were also measured. These differences are sensitive to NP effects in the $b \to s \ell^+ \ell^-$ transition and can deviate from zero if LFU breaking occurs. Their measurement is particularly interesting given the tensions with respect to the SM predictions observed in recent years in the P_5' observable of the $B^0 \to K^{*0} \mu^+ \mu^-$ decay [35] and in the branching fraction ratios between muons and electrons, like R_K [33] and R_{K^*} [34]. In the measurement presented in this thesis, a counting method approach derived from the symmetry properties of the *CP*-averaged differential decay width of the $B^0 \to K^{*0}\ \ell^+ \ell^-$ decay was used to determine the values of the $A_{\rm FB}$, S_4, S_5, S_7, and S_8 observables. This procedure has the advantage of being robust even with limited statistics and does not strongly rely on the knowledge of the angular distribution of the backgrounds. The measurement required a good understanding of the reconstruction effects of electrons and muons in the LHCb detector, which is helpful in view of future measurements of the same type. The angular observables were measured in two regions of q^2: in the low q^2 region from $0.1\,{\rm GeV}^2/c^4$ to $1.1\,{\rm GeV}^2/c^4$ and in the central q^2 region from $1.1\,{\rm GeV}^2/c^4$ to $7.0\,{\rm GeV}^2/c^4$. While previous measurements at LHCb [34] considered an upper bound of the central q^2 region of $6.0\,{\rm GeV}^2/c^4$, the adoption of a novel selection in this measurement allowed to extend the central q^2 region up to $7.0\,{\rm GeV}^2/c^4$, with a consequent significant increase in the statistics available. This was achieved by applying a combined selection in the normal q^2 and in the constrained q^2. The latter was obtained by constraining the mass of the B^0 candidate to the nominal mass of the B^0 meson. This additional requirement suppressed the contamination from $B^0 \to K^{*0} J/\psi\ (\to e^+ e^-)$ decays leaking in the central q^2 region. This study also showed that it is feasible to increase the lower bound of the low q^2 region from $0.045\,{\rm GeV}^2/c^4$, as in previous LHCb measurements [34], to $0.1\,{\rm GeV}^2/c^4$. This is preferred because it allows to easily compare measurements with muons and electrons in the final state, since both leptons can be considered massless above this threshold. The results of the measurement are compatible with the SM predictions and are dominated by the statistical uncertainties. For this reason, it would be advantageous to repeat the measurement including the LHC Run II data. The increased statistics will also allow to perform a full angular maximum likelihood fit of the $B^0 \to K^{*0} e^+ e^-$ decay, similarly to what done in Ref. [35] for the muon mode. In view of this measurement, the counting method approach would offer a strong cross-check, given its robustness. With the statistics collected in the LHC Run I and Run II, a measurement of the Q_i and D_i observables would become feasible at LHCb, thus allowing to extend our knowledge on potential NP effects that violate LFU and hopefully shading some light on the intriguing anomalies observed in recent years.

Acknowledgements

"No one undertakes research in physics with the intention of winning a prize. It is the joy of discovering something no one knew before." – Stephen Hawking

When I first moved to Zurich I was not sure what I would have found here, what my research topic would have been and which people I would have met and worked with. At the end of my adventure at the University of Zurich, nearly four years later, I can say that the PhD gave me the opportunity to meet incredible people and to work in an international and dynamic environment where I had the freedom to follow my research interests and grow as a scientist but also, and above all, as a person.

Somebody once told me that a great boss is hard to find, difficult to leave, and impossible to forget. I could not imagine a better sentence to describe my supervisor, Nico, who supported me during every moment of my PhD and was an inexhaustible source of fruitful ideas. Nico, you have always been able to ask the right questions at the right time to shade some light on the countless challenges that we faced in our daily work. Chatting with you has been a constant source of inspiration and a way to keep up the motivation even after spending days in debugging obscure pieces of code.

A sincere thank you goes to Ueli, Olaf, and Katharina, for their valuable help in a wide range of scientific and practical aspects of my life in Zurich. I remember when we went to the lab together, Ueli, and you were able to fix the issues related to the noise of our laser test stand in less than three minutes, explanation included. I am very grateful to you, Olaf, for teaching me most of the things I know about silicon sensors. Thanks for all the discussions we had and for your incredible attention to detail. Even when we did not share the same opinion, I appreciated your frankness and directness and I definitely have to thank you for always pushing me to do better. Thanks Katharina, I got so many good advices from you and your pragmatic approach was very helpful in keeping in mind the different priorities and in planning things ahead of the deadlines. Thanks for all the coffee breaks, the chats, and last but not least, the patience in speaking slowly and helping me with improving my German.

Thanks to you, Rafael and Albert, for revising my work and giving me precious comments and suggestions at any time of the day or night.

Michele and Andrea, I cannot forget the afternoons spent in discussing about the profound meaning of what we were doing and the tons of debugging and coding we made. I am sure you will have a lot of fun fitting invariant masses and angular distributions in the next months.

These years have been a continuos discovery and I cannot forget all the people that helped me in the early days in Zurich and never stopped to be a reference point for me. Achim, Arno, and Roman, I remember the first time you showed me how to use the soldering iron so that I could assemble my Arduino kit, and the great time we spent together in the lab in the first part of my PhD. A huge thanks goes to you, Stefan, for being such a friendly and generous person. I am glad we became friends and it is always a pleasure to meet you and Monika. Christian, Nicola, and Andreas, it was great to organise the python school together with you and I hope we will keep working together in the future. I have to thank you too, Roland, for every time I bothered you because of some issues with the farm in Zurich and for your ability to reply 24 hours a day and seven days a week. I promise I will free some space in my home area in the following weeks. Thanks to Carmelina,

Monika, Brigitte, and Ruth, for being so helpful and for allowing me to practice my German every time we discussed about something.

A special thanks goes to Marzia and Annarita, for being two such amazing friends. I do not know how you managed to have so much patience to wait for me every morning when I was late for our routine coffee. The (second) breakfast together and our random chats in front of a delicious Schoggi-Gipfeli and a not so delicious cappuccino were definitely the best way I could imagine to start a productive day at work. We even managed to become good friends with the ladies working in the cafeteria! And since a healty mind works alongside a healthy body, as the old Romans used to say, I must thank you, Annarita and Davide, for keeping up the motivation and convincing me to go to the gym very, very regularly during the last months, well done! Thanks to Iaroslava, who shared the experience of doing a test beam at CERN with me, and to all my colleagues in Zurich and at CERN, for providing such a great environment in which to carry out a PhD.

Thanks to the friends I met (or met again) in Zurich, and in particular to Marco, Natalie, Niccolò, Valentina, Francesco, Marianna, Claudia, and Alessandra, we had a great time together!

These acknowledgements would not be complete without a special thanks to my best friends from the University of Pisa, and in particular to Stefania, Adele, Laura, Anna, Alberto, Stefano, and Roberto. We are spread over several continents now but I am glad we always try to find a way to meet each other from time to time: it can be for a pizza in Meyrin, for a tagliata in Pisa, or for our celebrations of the New Year abroad, but it is always great to go out together as in the old days in Pisa. For not mentioning the random and often surreal chats on the most diverse topics in our WhatsApp groups.

Going back of few years, I want to thank you some of my old-time friends from high or even from primary school: Alessia, Francesca, Marianna, Ada Rita, Sara, Agostino, Rocco, and Salvatore. It is always cool to come back to Maglie and go out for a drink or an aperitif together, for not mentioning that you always manage to keep me up-to-date with the latest gossips of my home town.

No words can suffice to thank my family, who has always been there to share the ordinary and extraordinary happenings of my life. Despite how far we can be, I know we are very close to each other and I would not be the person I am if it were not for your constant and unconditional support. You are there every time I need an advice, but you also taught me to take the freedom and the responsibility of my choices.

And it is eventually time to thank you, Marco, for sharing your life with me and for making each and every minute we spend together a very special moment. Thanks for having prepared countless and delicious dinners when I was too busy with the writing of my thesis. After many years apart from each other I know how lucky I am now that I can share my daily life with you. Thanks for being able to understand what I have in mind even before I speak and for supporting me during all these years together. You are the best partner I could imagine.

Bibliography

[1] F. Pedregosa *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[2] https://github.com/yandex/rep.

[3] T. Likhomanenko *et al.* Reproducible Experiment Platform. *Journal of Physics: Conference Series*, 664(5):052022, 2015.

[4] https://arogozhnikov.github.io/hep_ml/.

[5] S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.*, 22:579–588, 1961.

[6] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[7] A. Salam. Weak and electromagnetic interactions. *Conf. Proc.*, C680519:367–377, 1968.

[8] Wikimedia Commons. The Standard Model of elementary particles. 2006. https://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg.

[9] F. Englert *et al.* Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.

[10] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[11] G. S. Guralnik *et al.* Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, 1964.

[12] R. Aaij *et al.* (LHCb collaboration). Observation of $J/\psi\, p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^- p$ decays. *Phys. Rev. Lett.*, 115:072001, 2015.

[13] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[14] M. Kobayashi *et al.* CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[15] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[16] E. K. Akhmedov. Neutrino physics. 1999. arXiv:hep-ph/0001264.

[17] B. Pontecorvo. Inverse beta processes and nonconservation of lepton charge. *Sov. Phys. JETP*, 7:172–173, 1958.

[18] Z. Maki *et al.* Remarks on the Unified Model of Elementary Particles. *Prog. Theor. Phys.*, 28:870–880, 1962.

[19] A. M. Baldini *et al.* (MEG collaboration). Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment. *Eur. Phys. J.*, C76(8):434, 2016.

[20] A. Bravar. The Mu3e Experiment. *Nuclear and Particle Physics Proceedings*, 287–288:169–172, 2017.

[21] R. M. Carey *et al.* (Mu2e collaboration). Proposal to search for $\mu^- N \to e^- N$ with a single event sensitivity below 10^{-16}. 2008. FERMILAB-PROPOSAL-0973.

[22] Y. Nakatsugawa (DeeMe collaboration). Search for Muon to Electron Conversion in Nuclear Field at J-PARC MLF. *Nuclear and Particle Physics Proceedings*, 273:1692–1698, 2016.

[23] Y. G. Cui *et al.* (COMET collaboration). Conceptual design report for experimental search for lepton flavor violating μ-e conversion at sensitivity of 10^{-16} with a slow-extracted bunched proton beam (COMET). 2009. KEK-2009-10.

[24] G. Buchalla *et al.* Weak decays beyond leading logarithms. *Rev. Mod. Phys.*, 68:1125–1244, 1996.

[25] A. Petermann. Normalization of constants in the quanta theory. *Helv. Phys. Acta*, 26:499–520, 1953.

[26] M. Gell-Mann *et al.* Quantum electrodynamics at small distances. *Phys. Rev.*, 95:1300–1312, 1954.

[27] A. J. Buras. Theoretical review of B-physics. *Nucl. Instrum. Methods Phys. Res., A*, 368(1):1 – 20, 1995.

[28] A. J. Buras. Operator product expansion, renormalization group and weak decays. *Lect. Notes Phys.*, 558:65–85, 2000.

[29] T. Blake *et al.* Rare b hadron decays at the LHC. *Annual Review of Nuclear and Particle Science*, 65(1):113–143, 2015.

[30] C. Bobeth *et al.* Angular distributions of $\bar{B} \to \bar{K}\bar{l}l$ decays. *JHEP*, 12:040, 2007. arXiv:0709.4174.

[31] D. Becirevic *et al.* Leptoquark model to explain the B-physics anomalies, R_K and R_D. *Phys. Rev. D*, 94:115021, 2016.

[32] P. Cox *et al.* SU(5) Unification with TeV-scale Leptoquarks. *JHEP*, 03:035, 2017.

[33] R. Aaij *et al.* (LHCb collaboration). Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 113:151601, 2014. arXiv:1406.6482.

[34] R. Aaij *et al.* (LHCb collaboration). Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays. *JHEP*, 08:055, 2017.

[35] R. Aaij *et al.* (LHCb collaboration). Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,\text{fb}^{-1}$ of integrated luminosity. *JHEP*, 02:104, 2016. arXiv:1512.04442.

[36] F. Kruger *et al.* Probing new physics via the transverse amplitudes of $B^0 \to K^{*0}(\to K^-\pi^+)l^+l^-$ at large recoil. *Phys. Rev.*, D71:094009, 2005. arXiv:hep-ph/0502060.

[37] S. Descotes-Genon *et al.* Implications from clean observables for the binned analysis of $B \to K^*\mu^+\mu^-$ at large recoil. *JHEP*, 01:048, 2013. arXiv:1207.2753.

[38] `https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/RareDecayConferenceMaterial`.

[39] R. Aaij *et al.* (LHCb collaboration). Angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay in the low-q^2 region. *JHEP*, 04:064, 2015. arXiv:1501.03038.

[40] D. Becirevic *et al.* On transverse asymmetries in $B \to K^*\ell^+\ell^-$. *Nucl. Phys.*, B854:321, 2012. arXiv:1106.3283.

[41] S. Jäger *et al.* Reassessing the discovery potential of the $B \to K^*\ell^+\ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities. *Phys. Rev. D*, 93:014028, 2016. arXiv:1412.3183.

[42] J. P. Lees *et al.* (BABAR collaboration). Measurement of branching fractions and rate asymmetries in the rare decays $b \to K^{(*)}\ell^+\ell^-$. *Phys. Rev. D*, 86:032012, 2012.

[43] J.-T. Wei *et al.* (Belle collaboration). Measurement of the differential branching fraction and forward-backward asymmetry for $b \to K^{(*)}l^+l^-$. *Phys. Rev. Lett.*, 103:171801, 2009.

[44] C. Bouchard *et al.* Standard Model predictions for $B \to K\ell^+\ell^-$ with form factors from lattice QCD. *Phys. Rev. Lett.*, 111:162002, 2013. arXiv:1306.0434.

[45] R. Gauld *et al.* An explicit $Z^{'}$-boson explanation of the $B \to K^*\mu^+\mu^-$ anomaly. *JHEP*, 01:069, 2014. arXiv:1310.1082.

[46] A. J. Buras *et al.* The anatomy of quark flavour observables in 331 models in the flavour precision era. *JHEP*, 02:023, 2013. arXiv:1211.1237.

[47] W. Altmannshofer *et al.* New physics in $B \to K^*\mu\mu$? *Eur. Phys. J.*, C73:2646, 2013. arXiv:1308.1501.

[48] R. Aaij *et al.* (LHCb collaboration). Measurements of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction. *JHEP*, 11:047, 2016. [Erratum: JHEP, 04:142 (2017)], arXiv:1606.04731.

[49] W. Altmannshofer *et al.* Interpreting hints for lepton flavor universality violation. *Phys. Rev.*, D96(5):055008, 2017. arXiv:1704.05435.

[50] R. Aaij *et al.* (LHCb collaboration). Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$. *Phys. Rev. Lett.*, 115(11):111803, 2015. [Erratum: Phys. Rev. Lett., 115:15,159901(2015)], arXiv:1506.08614.

[51] R. Aaij *et al.* (LHCb collaboration). Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays. 2017. arXiv:1708.08856.

[52] R. Aaij *et al.* (LHCb collaboration). Measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays. 2017. arXiv:1711.02505.

[53] A. Greljo *et al.* On the breaking of Lepton Flavor Universality in B decays. *JHEP*, 1507:142, 2015. arXiv:1506.01705.

[54] S. Wehle *et al.* (Belle collaboration). Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$. *Phys. Rev. Lett.*, 118(11):111801, 2017. arXiv:1612.05014.

[55] N. Serra *et al.* Measuring the breaking of lepton flavor universality in $B \to K^*\ell^+\ell^-$. *Phys. Rev.*, D95:035029, 2017.

[56] B. Capdevila *et al.* Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data. *JHEP*, 01:093, 2018. arXiv:hep-ph/1704.05340.

[57] C. Lefevre. The CERN accelerator complex. 2008. CERN-DI-0812015, https://cds.cern.ch/record/1260465.

[58] N. Cartiglia. Total pp cross section measurements at 2, 7, 8 TeV and 57 TeV. Technical Report CERN-CMS-CR-2013-059, 2013.

[59] R. Aaij *et al.* (LHCb collaboration). Measurement of $\sigma(pp \to b\bar{b}\mathrm{X})$ at $\sqrt{s} = 7$ TeV in the forward region. *Physics Letters B*, 694(3):209–216, 2010.

[60] R. Aaij *et al.* (LHCb collaboration). Prompt charm production in pp collisions at $\sqrt{s} = 7$ TeV. LHCb-CONF-2010-013, 2010.

[61] R. Aaij *et al.* (LHCb collaboration). Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions. *Phys. Rev. Lett.*, 118(5):052002, 2017. [Erratum: Phys. Rev. Lett., 119:16,169901 (2017)], arXiv:1612.05140.

[62] V. P. Andreev. B production at the LHC/QCD aspects. *arXiv:0706.1789*, 2007.

[63] M. L. Mangano. Two lectures on heavy quark production in hadronic collisions. Technical Report CERN-TH-97-328, 1997.

[64] http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[65] A. A. Alves Jr *et al.* (LHCb collaboration). The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, 2008.

[66] M. Lamont *et al.* (LHC collaboration). LHC report: level best. Dernières nouvelles du LHC: à son meilleur niveau. (BUL-NA-2012-345. 47/2012):4, 2012.

[67] R. Aaij *et al.* (LHCb collaboration). LHCb detector performance. *Int. J. Mod. Phys. A*, 30:1530022, 2015. arXiv:1412.6352.

[68] CERN. LHC Study Group. *LHCb: Technical Proposal*. Tech. Proposal. 1998.

[69] LHCb - Large Hadron Collider beauty experiment. http://lhcb-public.web.cern.ch/lhcb-public/.

[70] P. R. Barbosa-Marinho *et al.* (LHCb collaboration). LHCb VELO (VErtex LOcator): technical design report. Technical Report CERN-LHCC-2001-011, 2001.

[71] P. R. Pérez. The LHCb vertex locator performance and vertex locator upgrade. *arXiv:1209.4845*, 2012.

[72] R. P. Bernhard *et al.* The LHCb silicon tracker. Technical Report CERN-LHCb-2007-126, 2007.

[73] P. R. Barbosa-Marinho *et al.* (LHCb collaboration). LHCb outer tracker: technical design report. Technical Report CERN-LHCC-2001-024, 2001.

[74] Ph. d'Argent *et al.* Improved performance of the LHCb Outer Tracker in LHC Run 2. *Journal of Instrumentation*, 12:P11016, 2017.

[75] M. Adinolfi *et al.* Performance of the LHCb RICH detector at the LHC. *arXiv:1211.6759*, 2012.

[76] I. Machikhiliyan *et al.* (LHCb collaboration). The LHCb electromagnetic calorimeter. *Journal of Physics: Conference Series*, 160(1):12047, 2009.

[77] Y. Guz *et al.* (LHCb collaboration). The LHCb hadron calorimeter. *Journal of Physics: Conference Series*, 160(1):12054, 2009.

[78] S. Filippov *et al.* Experimental performance of SPD/PS detector prototypes. Technical Report LHCb-2000-031, 2000.

[79] A. A. Alves Jr *et al.* Performance of the LHCb muon system. *Journal of Instrumentation*, 8(02):P02022, 2013.

[80] R. Aaij *et al.* (LHCb collaboration). The LHCb trigger and its performance. *arXiv:1211.3055*, 2012.

[81] R. Aaij *et al.* (LHCb collaboration). Absolute luminosity measurements with the LHCb detector at the LHC. *arXiv:1110.2866*, 2011.

[82] E. Aslanides *et al.* The level–0 muon trigger for the LHCb experiment. *arXiv:0705.0310*, 2007.

[83] A. M. Sanchez. Performances of the LHCb L0 calorimeter trigger. Technical Report CERN-LHCb-PUB-2011-026, 2012.

[84] Trigger schemes. https://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html.

[85] V. V. Gligorov. A single track HLT1 trigger. Technical Report CERN-LHCb-PUB-2011-003, 2011.

[86] P. R. Barbosa-Marinho *et al.* (LHCb collaboration). LHCb online system technical design report: data acquisition and experiment control. Technical Report CERN-LHCC-2001-040, 2001.

[87] R. Jacobsson. Central FPGA–based destination and load control in the LHCb MHz event readout. *Nucl. Instrum. Methods Phys. Res., A*, 688:41–50, 2011.

[88] R. Aaij *et al.* (LHCb collaboration). Letter of intent for the LHCb upgrade. 2011. CERN-LHCC-2011-001.

[89] R. Aaij *et al.* (LHCb collaboration). Framework TDR for the LHCb upgrade: technical design report. 2012. CERN-LHCC-2012-007, LHCB-TDR-012.

[90] A. A. Alves Jr. *et al.* (LHCb collaboration). LHCb tracker upgrade technical design report. 2014. CERN-LHCC-2014-001, LHCB-TDR-015.

[91] R. Aaij *et al.* (LHCb collaboration). LHCb VELO upgrade technical design report. 2013. CERN-LHCC-2013-021, LHCB-TDR-013.

[92] M. van Beuzekom *et al.* VeloPix ASIC development for LHCb VELO upgrade. *Nucl. Instrum. Methods Phys. Res., A*, 731:92 – 96, 2013.

[93] R. Aaij *et al.* (LHCb collaboration). LHCb particle identification upgrade technical design report. 2013. CERN-LHCC-2013-022, LHCB-TDR-014.

[94] R. Aaij *et al.* (LHCb collaboration). LHCb trigger and online upgrade technical design report. 2014. CERN-LHCC-2014-016, LHCB-TDR-016.

[95] A. A. Alves Jr. *et al.* (LHCb collaboration). The LHCb detector at the LHC. *Journal of Instrumentation*, 3:S08005, 2008.

[96] A. A. Alves Jr. *et al.* (LHCb collaboration). LHCb trigger and online upgrade technical design report. 2014. CERN-LHCC-2014-016, LHCB-TDR-016.

[97] E. Bowen *et al.* VeloUT tracking for the LHCb upgrade. 2013. LHCb-PUB-2013-023, CERN-LHCb-PUB-2013-023.

[98] T. T. Bohlen *et al.* The FLUKA code: developments and challenges for high energy and medical applications. *Nuclear Data Sheets*, 120:211–214, 2014.

[99] A. Ferrari *et al.* FLUKA: a multi-particle transport code. 2005. CERN-2005-10.

[100] P. Moreira. The GBTx link interface ASIC. 2012. https://espace.cern.ch/GBT-Project/GBTX/Specifications/gbtxSpecsV1.8.pdf.

[101] C. Grupen *et al.* Particle detectors. *Cambridge University Press*, 2008.

[102] H. Spieler. Semiconductor detector systems. *Oxford University Press*, 2005.

[103] M. Moll. Radiation damage in silicon particle detectors: microscopic defects and macroscopic properties. *PhD thesis, University of Hamburg*. DESY-THESIS-1999-040.

[104] G. Casse *et al.* Charge collection and charge sharing in heavily irradiated n-side read-out silicon microstrip detectors. *Nucl. Instrum. Meth. A*, 511:112–117, 2003.

[105] G. Casse *et al.* Improving the radiation hardness properties of silicon detectors using oxygenated n-type and p-type silicon. *IEEE Trans. Nucl. Sci.*, 47, 1:527–532, 2000.

[106] M. Artuso *et al.* Signal coupling to embedded pitch adapters in silicon sensors. 2017. LHCb-PUB-2017-017, CERN-LHCb-PUB-2017-017.

[107] Hamamatsu Photonics. http://www.hamamatsu.com.

[108] A. Abba *et al.* Testbeam studies of pre-prototype silicon strip sensors for the LHCb UT upgrade project. *Nucl. Instrum. Meth. A*, 806:244–257, 2016.

[109] S. Löchner and M. Schmelling. The Beetle Reference Manual - chip version 1.3, 1.4 and 1.5. 2006. LHCb-2005-105, CERN-LHCb-2005-105, http://www.kip.uni-heidelberg.de/lhcb/Publications/BeetleRefMan_v1_3.pdf.

[110] Nuclear Instruments srl, IT-22045 Lumbrugo (CO), Italy. http://www.nuclearinstruments.eu/.

[111] A. Abba *et al.* Studies of pre-prototype sensors for the UT upgrade project. 2015. LHCb-PUB-2015-006.

[112] K. Akiba *et al.* The Timepix telescope for high performance particle tracking. *Nucl. Instrum. Meth. A*, 723:47, 2013. arXiv:1304.5175.

[113] E. Frojdh *et al.* Timepix3: first measurements and characterization of a hybrid-pixel detector working in event driven mode. *Journal of Instrumentation*, 10(01):C01039, 2015.

[114] J. Visser *et al.* SPIDR: a read-out system for Medipix3 and Timepix3. *Journal of Instrumentation*, 10:C12028, 2015.

[115] F. James. MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1. 1994. CERN-D-506.

[116] V. Blobel. Software alignment for tracking detectors. *Nucl. Instrum. Meth. A*, 566:S5–S13, 2006.

[117] A. Abba *et al.* Study of prototype sensors for the Upstream Tracker Upgrade. LHCb-PUB-2016-007.

[118] G. Casse *et al.* Degradation of charge sharing after neutron irradiation in strip silicon detectors with different geometries. *Nucl. Instrum. Methods Phys. Res. A*, 730:54–57, 2013. AIDA-PUB-2013-032.

[119] Synopsys. http://www.synopsys.com.

[120] E. H. Nicollian *et al.* MOS (Metal Oxide Semiconductor) Physics and Technology. *John Wiley and Sons, New York*, 1982.

[121] D. van Dyk *et al.* EOS – A HEP Program for Flavour Observables, 2017. `https://eos.github.io/`.

[122] A. Bharucha *et al.* $B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules. *JHEP*, 2016(8):98, 2016.

[123] A. Ali *et al.* Comparative study of the decays $\overrightarrow{B}$ $(K, K^*)l^+l^-$ in the standard model and supersymmetric theories. *Phys. Rev. D*, 61:074024, 2000.

[124] R. Aaij *et al.* Muon triggers in the High Level Trigger of LHCb. Technical Report LHCb-PUB-2011-017. CERN-LHCb-PUB-2011-017, 2011.

[125] A. Poluektov. Kernel density estimation of a multidimensional efficiency profile. *Journal of Instrumentation*, 10(02):P02011, 2015. arXiv:1411.5528.

[126] A. Rogozhnikov. Reweighting with boosted decision trees. *Journal of Physics: Conference Series*, 762:012036, 2016. arXiv:1608.05806.

[127] M. Pivk *et al.* sPlot: A statistical tool to unfold data distributions. *Nucl. Instrum. Meth. A*, 555:356–369, 2005. arXiv:physics/0402083.

[128] T. M. Karbach *et al.* Decay time integrals in neutral meson mixing and their efficient evaluation. 2014. arXiv:1407.0748.

[129] Jonas Eschle. Studies of $B^+ \to K^+\pi^-\pi^+e^+e^-$ decays at LHCb. *Bachelor thesis, University of Zurich*, 2017. CERN-THESIS-2015-117.

[130] R. Aaji *et al.* (LHCb collaboration). Measurement of b hadron production fractions in 7 TeV pp collisions. *Phys. Rev. D*, 85:032008, 2012.

[131] R. Aaij *et al.* (LHCb collaboration). Measurement of the $B^0 \to K^{*0}e^+e^-$ branching fraction at low dilepton mass. *JHEP*, 05:159, 2013. arXiv:1304.3035.

[132] R. Aaij *et al.* (LHCb collaboration). Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics. *JHEP*, 04:001, 2013. arXiv:1301.5286.

[133] M. Bordone *et al.* On the standard model predictions for R_K and R_{K^*}. *Eur. Phys. J.*, C76(8):440, 2016.

[134] R. Aaij *et al.* (LHCb collaboration). Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays. *Phys. Rev. D*, 88:052002, 2013.

[135] F. James. Minuit: function minimization and error analysis. `https://root.cern.ch/download/minuit.pdf`.

[136] J. Gratrex *et al.* Generalized helicity formalism, higher moments, and the $B \to K_{J_K}(\to K\pi)\overline{\ell_1}\ell_2$ angular distributions. *Phys. Rev. D*, 93:054008, 2016.

[137] B. Efron. An introduction to the bootstrap. 1994.

APPENDIX A

Uniform classifiers

This appendix provides some additional details on the uniform classifiers described in Chapter 6. In particular, the ROC curve of the uniform classifiers is shown in Fig. A.1, while Fig. A.2 shows the classifier response for the training and testing samples of signal and background. The efficiency as a function of the uniform variables is shown in Fig. A.3 for the classifiers with uniformity along the B^0 invariant mass only and in Fig. A.4 for the classifiers with uniformity along the B^0 invariant mass and the decay angles θ_ℓ, θ_K, and ϕ.

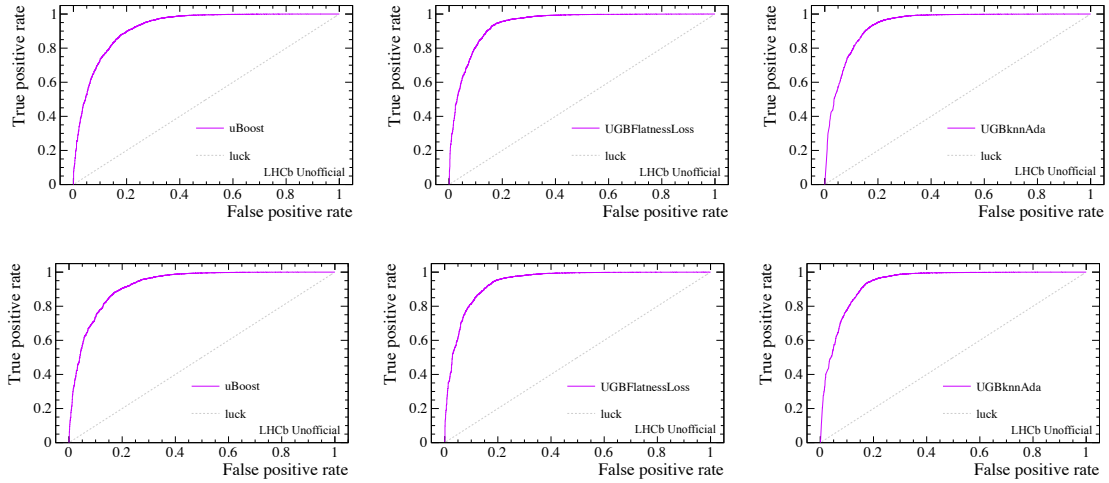


Figure A.1: ROC curve of the `uBoost` (left), `UGBknnAda` (center), and `UGBFlatnessLoss` (right) uniform classifiers with uniformity along the B^0 invariant mass only (top) and along the B^0 invariant mass and the decay angles θ_ℓ, θ_K, and ϕ (bottom).

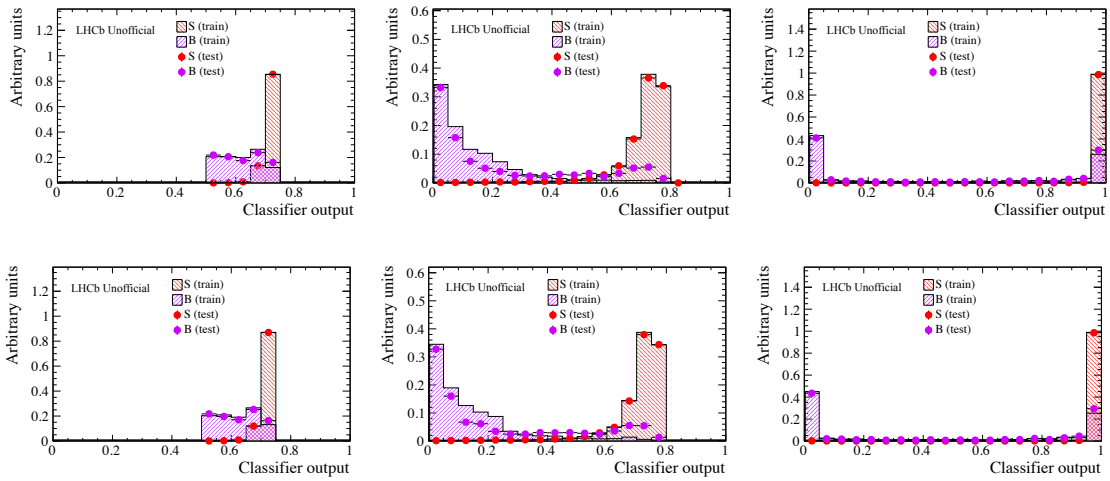


Figure A.2: Classifier response for the training and testing samples of signal and background for the `uBoost` (left), `UGBknnAda` (center), and `UGBFlatnessLoss` (right) uniform classifiers with uniformity along the B^0 invariant mass (top) and along the B^0 invariant mass and the decay angles θ_ℓ, θ_K, and ϕ (bottom).

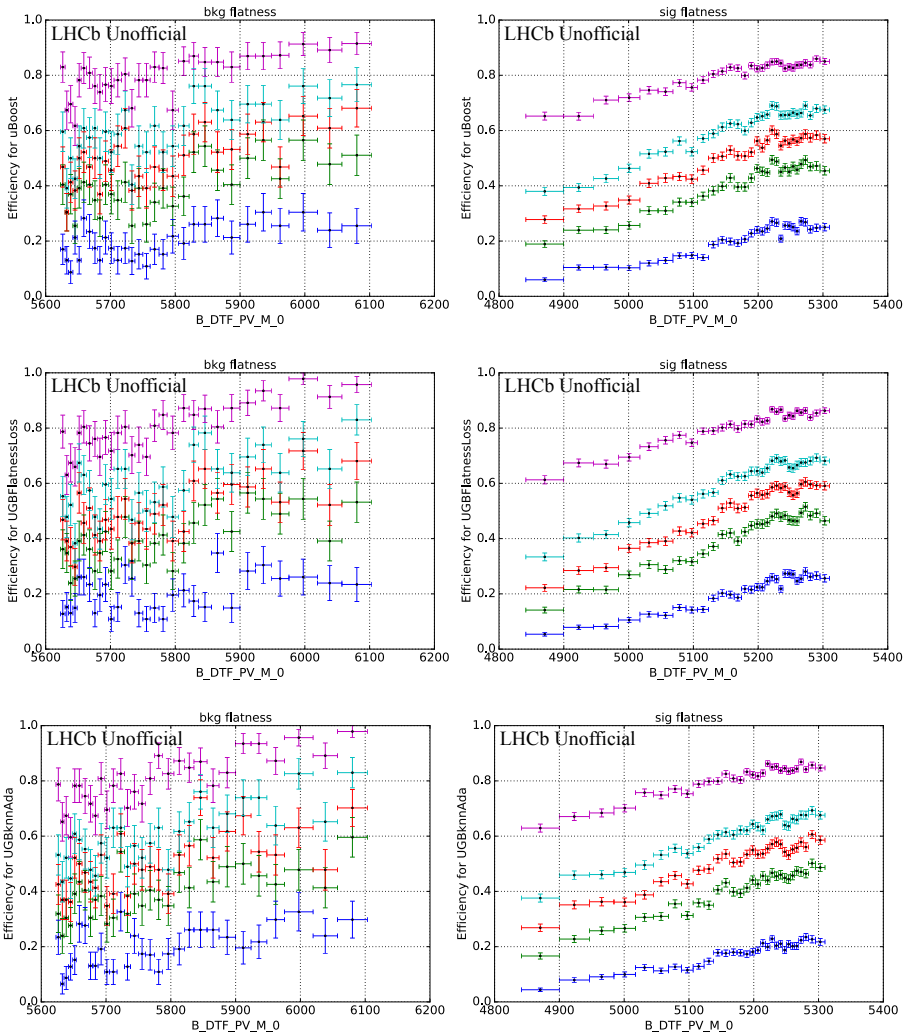


Figure A.3: Efficiency as a function of the B^0 invariant mass on the background (left) and signal (right) samples for the `uBoost` (top), `UGBknnAda` (middle), and `UGBFlatnessLoss` (bottom) uniform classifiers with uniformity along the B^0 invariant mass.

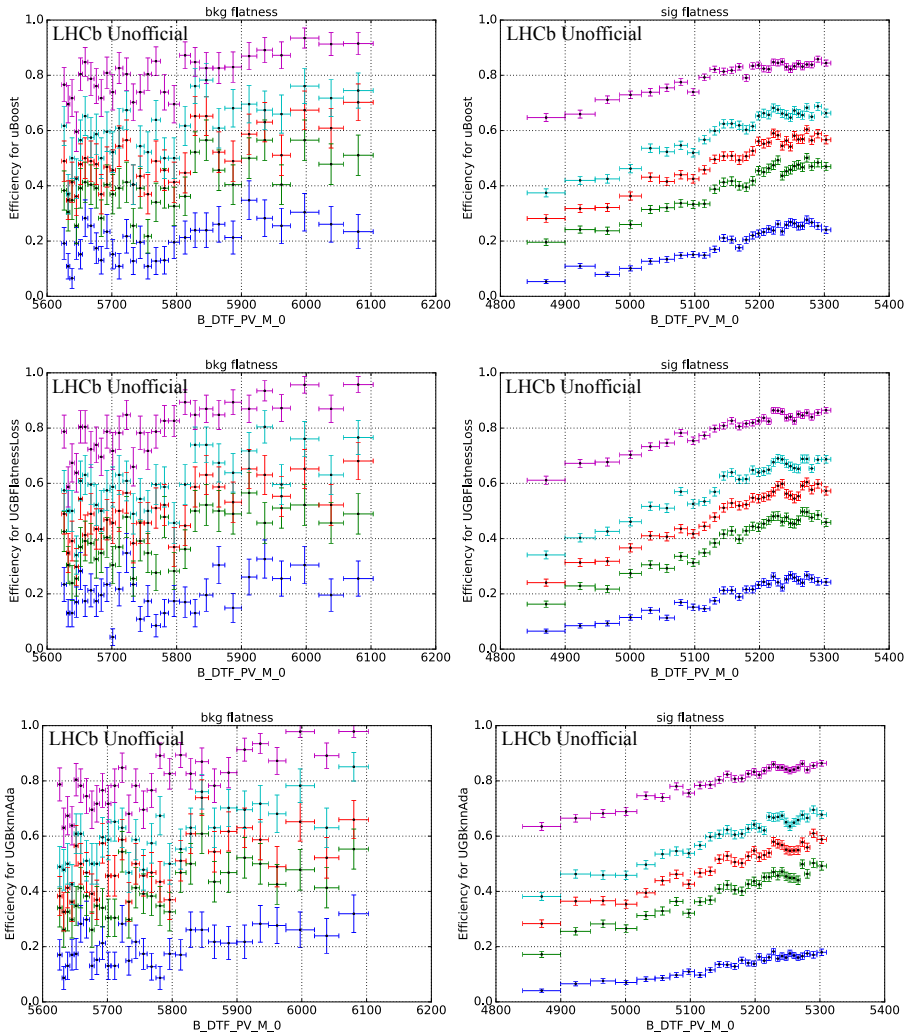


Figure A.4: Efficiency as a function of the B^0 invariant mass on the background (left) and signal (right) samples for the `uBoost` (top), `UGBknnAda` (middle), and `UGBFlatnessLoss` (bottom) uniform classifiers with uniformity along the B^0 invariant mass and the decay angles θ_ℓ, θ_K, and ϕ.